16869K-109000US

日本国特許庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日
Date of Application: ２００３年１０月　１日

出願番号
Application Number: 特願２００３－３４３４７８

[ST. 10/C]: [ＪＰ２００３－３４３４７８]

出願人
Applicant(s): 株式会社日立製作所

２００３年１１月１２日

特許庁長官
Commissioner,
Japan Patent Office 今井康夫

出証番号　出証特２００３－３０９３６１７

【書類名】　特許願
【整理番号】　340301000
【提出日】　平成15年10月 1日
【あて先】　特許庁長官殿
【国際特許分類】　G06F 12/00
【発明者】
　【住所又は居所】　神奈川県川崎市麻生区王禅寺１０９９番地　株式会社日立製作所システム開発研究所内
　【氏名】　富田　亜紀
【特許出願人】
　【識別番号】　000005108
　【氏名又は名称】　株式会社日立製作所
【代理人】
　【識別番号】　110000176
　【氏名又は名称】　一色国際特許業務法人
　【代表者】　一色　健輔
【手数料の表示】
　【予納台帳番号】　211868
　【納付金額】　21,000円
【提出物件の目録】
　【物件名】　特許請求の範囲　1
　【物件名】　明細書　1
　【物件名】　図面　1
　【物件名】　要約書　1

【書類名】特許請求の範囲
【請求項１】
データの記憶領域を供給する記憶装置と、
外部からのアクセス要求を受け付けて、前記アクセス要求に応じて前記記憶領域に対するデータの読み出し／書き込みを行うアクセス処理部と、
前記記憶領域を用いて構成される論理的な記憶領域である論理ボリュームを管理する論理ボリューム管理部と、
本番系の業務に適用されている前記論理ボリュームである本番ボリュームと、前記本番ボリュームに書き込まれるデータの複製が書き込まれる前記論理ボリュームである複製ボリュームとを管理するボリューム管理部と、
前記複製ボリュームに障害が生じている場合に、当該複製ボリュームとは異なる他の複製ボリュームに書き込まれているデータを用いて、前記障害の内容に応じた方法により、当該複製ボリュームを復元する複製ボリューム復元部と、
を備えることを特徴とするデータＩ／Ｏ装置。
【請求項２】
請求項１に記載のデータＩ／Ｏ装置であって、
前記複製ボリュームのうちの少なくとも一つを、前記外部からのアクセス要求に対して読み出し専用の複製ボリュームとして制御するアクセス制御部を有し、
前記複製ボリューム復元部は、前記読み出し専用に制御されている前記複製ボリュームに書き込まれているデータを前記障害が生じている複製ボリュームに複製することにより前記複製ボリュームを復元すること、
を特徴とするデータＩ／Ｏ装置。
【請求項３】
請求項１に記載のデータＩ／Ｏ装置であって、
前記複製ボリュームのうちの少なくとも一つを、前記外部からのアクセス要求に対して読み出し専用の複製ボリュームとして制御するアクセス制御部を有し、
前記複製ボリューム復元部は、前記障害が前記記憶装置のハードウエアに関する障害である場合に、前記読み出し専用に制御されている複製ボリュームを、前記障害が生じている前記複製ボリュームとして用いることにより前記複製ボリュームを復元すること、
を特徴とするデータＩ／Ｏ装置。
【請求項４】
請求項１に記載のデータＩ／Ｏ装置であって、
前記複製ボリュームのうちの少なくとも一つを、前記外部からのアクセス要求に対して読み出し専用の複製ボリュームとして制御するアクセス制御部を有し、
前記読み出し専用ボリュームに対するアクセス頻度を監視するアクセス頻度監視部を備え、
前記複製ボリューム復元部は、前記障害の生じている複製ボリュームに、前記読み出し専用に制御されている複製ボリュームのうち前記アクセス頻度が最小の複製ボリュームに書き込まれているデータを前記障害の生じている複製ボリュームに複製することにより前記障害が生じている複製ボリュームを復元すること、
を特徴とするデータＩ／Ｏ装置。
【請求項５】
請求項１に記載のデータＩ／Ｏ装置であって、
前記複製ボリュームのうちの少なくとも一つを、前記外部からのアクセス要求に対して読み出し専用の複製ボリュームとして制御するアクセス制御部を有し、
前記読み出し専用に制御されている複製ボリュームに対するアクセス頻度を監視するアクセス頻度監視部を備え、
前記複製ボリューム復元部は、前記障害の内容が前記記憶装置のハードウエアに関する障害である場合に、前記障害の生じている複製ボリュームに、前記読み出し専用に制御されている複製ボリュームのうち前記アクセス頻度が最小の複製ボリュームを、障害が生じ

ている前記複製ボリュームとして用いることにより前記障害が生じている複製ボリュームを復元すること、
　を特徴とするデータＩ／Ｏ装置。
【請求項６】
　請求項１に記載のデータＩ／Ｏ装置であって、
　前記複製ボリュームのうちの少なくとも一つを、前記外部からのアクセス要求に対して読み出し専用の複製ボリュームとして制御し、かつ、前記複製ボリュームのうちの少なくとも一つを、前記外部からのアクセス要求に対して書き込みを許可する複製ボリュームとして制御するアクセス制御部を有し、
　ある時刻以降に前記複製ボリュームに対して書き込まれたデータを前記論理ボリューム（差分ボリューム）に記憶する差分データ管理部を有し、
　前記複製ボリューム復元部は、前記書き込みを許可するように制御されている複製ボリュームに障害が生じている場合に、前記読み出し専用に制御されている複製ボリュームを前記差分ボリュームに書き込まれているデータを用いて更新することにより前記障害が生じている複製ボリュームを復元すること、
　を特徴とするデータＩ／Ｏ装置。
【請求項７】
　請求項１に記載のデータＩ／Ｏ装置であって、
　前記複製ボリュームのうちの少なくとも一つを、前記外部からのアクセス要求に対して読み出し専用の複製ボリュームとして制御し、かつ、前記複製ボリュームのうちの少なくとも一つを、前記外部からのアクセス要求に対して書き込みを許可する複製ボリュームとして制御するアクセス制御部を有し、
　前記読み出し専用に制御されている複製ボリュームに対するアクセス頻度を監視するアクセス頻度監視部を有し、
　ある時刻以降に前記複製ボリュームに対して行われた書き込みデータを前記論理ボリューム（差分ボリューム）に記憶する差分データ管理部を有し、
　前記複製ボリューム復元部は、前記書き込みを許可するように制御されている複製ボリュームに障害が生じている場合に、前記読み出し専用に制御されている複製ボリュームのうち前記アクセス頻度が最小の複製ボリュームを、前記差分ボリュームに書き込まれているデータを用いて更新することにより前記障害が生じている複製ボリュームを復元すること、
　を特徴とするデータＩ／Ｏ装置。
【請求項８】
　請求項２～７のいずれかに記載のデータＩ／Ｏ装置であって、
　ある時刻における前記複製ボリュームの内容に維持され、かつ、前記外部からのアクセス要求に対するデータの読み出し及び書き込みを禁止するように制御される論理ボリューム（スペアボリューム）を管理するスペアボリューム管理部を有し、
　前記読み出し専用のボリュームに代えて前記スペアボリュームを用いて前記障害が生じている複製ボリュームを復元すること、
　を特徴とするデータＩ／Ｏ装置。
【請求項９】
　請求項２～７のいずれかに記載のデータＩ／Ｏ装置であって、
　前記記憶装置を複数備え、
　前記複製ボリューム復元部は、前記障害の内容が前記記憶装置のハードウエアに関する障害である場合に、障害が生じている前記複製ボリュームを構成している前記記憶領域を、正常に動作している他の記憶装置を用いて構成し直すこと、
　を特徴とするデータＩ／Ｏ装置。
【請求項１０】
　請求項８に記載のデータＩ／Ｏ装置であって、
　前記記憶装置を複数備え、

前記複製ボリューム復元部は、前記障害の内容が前記記憶装置のハードウエアに関する障害である場合に、障害が生じている前記複製ボリュームを構成している前記記憶領域を、正常に動作している他の記憶装置を用いて構成し直すこと、
を特徴とするデータＩ／Ｏ装置。
【請求項１１】
請求項１に記載のデータＩ／Ｏ装置であって、
前記記憶装置はディスクドライブであり、
前記外部からのアクセス要求を送出してくる情報処理装置と通信する通信制御部を有することを特徴とするデータＩ／Ｏ装置。
【請求項１２】
データの記憶領域を供給する記憶装置と、
外部からアクセス要求を受け付けて、前記アクセス要求に応じて前記記憶領域に対するデータの読み出し／書き込みを行うアクセス処理部と、
前記記憶領域を用いて構成される論理的な記憶領域である論理ボリュームを管理する論理ボリューム管理部と、
を備えるデータＩ／Ｏ装置の制御方法であって、
前記データＩ／Ｏ装置は、
本番系の業務に適用されている前記論理ボリュームである本番ボリュームと、前記本番ボリュームに書き込まれるデータの複製が書き込まれる前記論理ボリュームである複製ボリュームとを管理し、
前記複製ボリュームの一つに障害が生じている場合に、当該複製ボリュームとは異なる他の複製ボリュームに書き込まれているデータを用いて、前記障害の内容に応じた方法により、当該複製ボリュームを復元すること、
を特徴とするデータＩ／Ｏ装置の制御方法。

【書類名】明細書
【発明の名称】データＩ／Ｏ装置及びデータＩ／Ｏ装置の制御方法
【技術分野】
【０００１】
本発明はデータＩ／Ｏ装置及びデータＩ／Ｏ装置の制御方法に関し、特に本番ボリュームに書き込まれるデータの複製が書き込まれる複製ボリュームについての可用性を確保するための技術に関する。
【背景技術】
【０００２】
膨大なデータを効率よく記憶し管理するためのストレージシステムは、近年、急速な発展を遂げているＩＴ関連作業のインフラとして、非常に重要な役割を担うようになってきている。また、このようにストレージに対する社会的な需要が高まる中で、ストレージシステムには、３６５日２４時間ノンストップ（無停止）での安定稼働が可能といった、非常に高い可用性が求められるようになってきている。このため、昨今のストレージシステムには、本番系業務に影響を与えずにデータをバックアップする仕組みやデータ分析や開発／テストのためにデータを複製する仕組み（レプリケーション技術）等、本番系業務の可用性を向上させるための様々な技術が取り入れられている（例えば、特許文献１を参照）。また、前記レプリケーション技術においては、本番系に適用されるボリューム（本番ボリューム）に記憶されているデータが他のボリューム（複製ボリューム）に複製されるが、この複製ボリュームは、データのバックアップやデータ分析、開発／テスト等、様々な副業務に利用される。つまり、これら副業務により本番系に与える影響が最小限に抑えられ、これによっても本番系業務の可用性の向上が図られている。
【特許文献１】米国特許第６，１０１，４９７号明細書
【発明の開示】
【発明が解決しようとする課題】
【０００３】
ところで、上記レプリケーション技術では、基本的に本番系業務で用いられる本番ボリュームの可用性の向上は図られるが、複製ボリュームの可用性については配慮されていない。しかしながら、現実の業務では、副業務で用いられる複製ボリュームについての可用性が求められることも少なくない。例えば、データ分析や開発、テスト等の業務において、プログラムに内在するバグやハードウエア障害等により、複製ボリュームの内容が破損することがあるが、そのような場合には、複製ボリュームを簡便かつ迅速に復元する仕組みが必要である。なお、複製ボリュームの復元に際し、本番ボリュームのデータを複製ボリュームに複製したとしても複製ボリュームの内容は必ずしも期待する内容に復元されるわけではない。複製時点における本番ボリュームの内容は既に更新されている可能性があり、復元後の内容が必ずしも破損前の複製ボリュームの内容に一致するとは限らないからである。
【０００４】
本発明はこのような背景に鑑みてなされたもので、本番ボリュームに書き込まれるデータの複製が書き込まれる複製ボリュームについての可用性を確保することが可能なデータＩ／Ｏ装置及びデータＩ／Ｏ装置の制御方法を提供することを目的とする。
【課題を解決するための手段】
【０００５】
上記目的を達成するための本発明のうち主たる発明は、データＩ／Ｏ装置であって、データの記憶領域を供給する記憶装置と、外部からのアクセス要求を受け付けて、前記アクセス要求に応じて前記記憶領域に対するデータの読み出し／書き込みを行うアクセス処理部と、前記記憶領域を用いて構成される論理的な記憶領域である論理ボリュームを管理する論理ボリューム管理部と、本番系の業務に適用されている前記論理ボリュームである本番ボリュームと、前記本番ボリュームに書き込まれるデータの複製が書き込まれる前記論理ボリュームである複製ボリュームとを管理するボリューム管理部と、前記複製ボリュー

ムに障害が生じている場合に、当該複製ボリュームとは異なる他の複製ボリュームに書き込まれているデータを用いて、前記障害の内容に応じた方法により、当該複製ボリュームを復元する複製ボリューム復元部と、を備えることとする。

【０００６】

前記記憶装置は、例えば、ディスクドライブ（ハードディスク装置）である。データＩ／Ｏ装置とは、例えば、情報処理装置から送信されるアクセス要求を受け付けて、前記アクセス要求に応じて前記ディスクドライブに対するデータの書き込みや読み出しを行うディスクアレイ装置である。本発明のＩ／Ｏデータ装置は、複製ボリュームの一つに障害が生じている場合に、当該複製ボリュームとは異なる他の複製ボリュームに書き込まれているデータを用いて、前記障害の内容に応じた方法により、当該複製ボリュームを復元する。本発明によれば、本番ボリュームに書き込まれるデータの複製が書き込まれる複製ボリュームの可用性を確保することができる。

【０００７】

また他の複製ボリュームや後述するスペアボリュームを用いて複製ボリュームを復元するようにしているため、例えば、データ分析や開発、テスト等において必要とされる必要な時点（ポジション）における複製ボリュームの内容の複製ボリュームを復元することが可能である。

【０００８】

複製ボリュームの復元は必ずしも一律の方法で行われるわけではなく、障害の内容に応じた方法により行われるので、効率よく柔軟な方法で複製ボリュームを復元することができる。障害の内容としては、例えば、データがソフトウエア的に破損している障害であるデータ障害、ディスクドライブのハードウエア故障に起因する障害であるハードウエア障害がある。復元方法としては、障害の生じている複製ボリュームの属性（読み出し専用や書き込み許可等）に応じた各種の方法があり、読み出し専用に制御されている前記複製ボリュームに書き込まれているデータを前記障害が生じている複製ボリュームに複製する方法、読み出し専用に制御されている複製ボリュームを前記障害が生じている前記複製ボリュームとして用いる方法、ある時刻以降に前記複製ボリュームに対して書き込まれたデータを前記論理ボリューム（差分ボリューム）に記憶しておき、読み出し専用に制御されている複製ボリュームを前記差分ボリュームに書き込まれているデータを用いて更新することにより書き込みを許可するように制御されている複製ボリュームを復元する方法、等がある。

【０００９】

上記復元に際して使用する読み出し専用ボリュームとしてアクセス頻度が最小のものを用いることにより、復元に関する処理が他の複製ボリュームを利用して行われている業務に与える影響を抑えることができ、複製ボリュームについて可用性を確保することができる。さらに上記復元に際して使用する読み出し専用ボリュームに代えて、ある時刻における前記複製ボリュームの内容に維持され、かつ、前記外部からのアクセス要求に対するデータの読み出し及び書き込みを禁止するように制御される論理ボリューム（スペアボリューム）を用いることにより、復元に関する処理が他の複製ボリュームを利用して行われている業務に与える影響を少なくすることができ、複製ボリュームについて高い可用性を確保することができる。さらにハードウエアに関する障害である場合には、障害が生じている前記複製ボリュームを構成している前記記憶領域を、正常に動作している他の記憶装置を用いて構成し直す。これにより複製ボリュームの識別子（例えば、論理ボリュームのＩＤ（ＬＩＤ））を変更することなく複製ボリュームを復元することができる。

【発明の効果】

【００１０】

本発明によれば、複製ボリュームについての可用性を確保することができる。

【発明を実施するための最良の形態】

【００１１】

以下、本発明の実施の形態について図面を用いて詳細に説明する。

【００１２】
＝＝＝ハードウエア構成＝＝＝
図１に本発明の一実施形態として説明するストレージシステムのハードウエア構成を示している。ストレージシステムは、情報処理装置としてのサーバ１００（１００－１、１００－２）、データＩ／Ｏ装置としてのディスク制御装置２００、管理サーバ１１０、等を備えて構成される。サーバ１００（１００－１、１００－２）、ディスク制御装置２００、管理サーバ１１０は、互いに通信可能に接続されている。前記通信の物理的なプロトコルとしては、例えば、イーサネット（登録商標）が用いられる。
【００１３】
ディスク制御装置２００は、サーバ１００（１００－１、１００－２）と接続され、サーバ１００（１００－１、１００－２）から送信されてくるデータの読み出し／書き込み要求（Read/Write要求）を受信する。なお、データの読み出し／書き込み要求はデータ入出力要求とも称する。ディスク制御装置２００は、記憶装置としてディスクドライブ２４０（２４０－１～２４０－５）を多数備えている。ディスク制御装置２００は、サーバ１００（１００－１、１００－２）から送信されてくるデータの入出力要求（アクセス要求）に応じてディスクドライブ２４０（２４０－１～２４０－５）に対するデータの読み出し／書き込みを行う。
【００１４】
ディスクドライブ２４０（２４０－１～２４０－５）は、サーバ１００（１００－１、１００－２）に提供するための物理的な記憶領域（以後、物理ボリュームと称する）を供給する。ディスク制御装置２００は、物理ボリュームを用いて構成される論理的な記憶領域である論理ボリューム（以降の説明ではＬＵ（Logical Unit）と称することもある）を単位として記憶領域を管理している。例えば、サーバ１００（１００－１、１００－２）は、論理ボリュームを指定することによりデータの書き込みもしくは読み出しの対象となるディスクドライブ２４０（２４０－１～２４０－５）上の記憶領域を特定することができる。なお、ディスクドライブ２４０は、図１に示すようにディスク制御装置２００と一体的になっている（例えば、ディスク制御装置２００が収容されている筐体と同一の筐体に収容される場合）こともあるし、ディスク制御装置２００とは別体となっている（例えば、ディスク制御装置２００が収容されている筐体とは別の筐体に収容される場合）こともある。
【００１５】
サーバ１００（１００－１、１００－２）は、ＣＰＵ（Central Processing Unit）、メモリ、入出力装置等を備えたコンピュータである。サーバ１００（１００－１、１００－２）はアクセスしてくる他のコンピュータに対して各種のサービスを提供する。前記サービスには、例えば、銀行の自動預金預け払いサービスやインターネットのホームページ閲覧サービスのようなオンラインサービス、科学技術分野における実験シミュレーションを行うバッチ処理サービス等がある。
【００１６】
サーバ１００（１００－１、１００－２）とディスク制御装置２００との間の通信は様々な通信プロトコルに従って行うようにすることができる。例えば、ファイバチャネルやＳＣＳＩ（Small Computer System Interface）、ＦＩＣＯＮ（Fibre Connection）（登録商標）、ＥＳＣＯＮ(Enterprise System Connection)（登録商標）、ＡＣＯＮＡＲＣ（Advanced Connection Architecture）（登録商標）、ＦＩＢＡＲＣ（Fibre Connection Architecture）（登録商標）、ＴＣＰ／ＩＰ（Transmission Control Protocol/Internet Protocol）等である。上記通信プロトコルを混在させるようにすることもできる。例えば、サーバ１００（１００－１、１００－２）がメインフレーム系のコンピュータである場合には、ＦＩＣＯＮやＥＳＣＯＮ、ＡＣＯＮＡＲＣ、ＦＩＢＡＲＣが用いられる。サーバ１００（１００－１、１００－２）がオープン系のコンピュータである場合には、例えば、ファイバチャネルやＳＣＳＩ、ＴＣＰ／ＩＰが用いられる。
【００１７】

サーバ１００（１００－１、１００－２）からのデータの読み出し／書き込み要求は、論理ボリュームにおけるデータの管理単位であるブロックを単位として行うようにすることもできるし、ファイル名を指定することによりファイル単位に行うようにすることもできる。後者の場合、ディスク制御装置２００は、サーバ１００（１００－１、１００－２）からのファイルレベルでのアクセスを実現するＮＡＳ（Network Attached Storage）として機能させることもできる。

【００１８】

ディスク制御装置２００は、磁気ディスク装置としてのディスクドライブ２４０（２４０－１～２４０－５）、サーバ１００（１００－１、１００－２）との間で通信を行う機能を提供するホストアダプタ（HA（Host Adaptor））２１０（２１０－１、２１０－２）、ディスクドライブ２４０（２４０－１～２４０－５）との間で通信を行う機能を提供するストレージアダプタ（SA）２３０、管理アダプタ（MA（Management Adaptor））２２０、内部ネットワーク２５０、等を構成要素として備えている。このうちホストアダプタ（HA）２１０（２１０－１、２１０－２）は、チャネルアダプタと称されることがあり、またストレージアダプタ（SA）２３０はディスクアダプタと称されることがある。内部ネットワーク２５０は、ホストアダプタ（HA）２１０（２１０－１、２１０－２）、ストレージアダプタ（SA（Storage Adaptor））２３０、管理アダプタ（MA）２２０、を互いに通信可能に接続する。内部ネットワーク２５０は、例えば、高速クロスバスイッチ等を用いて構成される。内部ネットワーク２５０には、ホストアダプタ（HA）２１０（２１０－１、２１０－２）とストレージアダプタ（SA）２３０との間で授受されるバッファとなるキャッシュメモリが接続されていることもある。ホストアダプタ（HA）２１０（２１０－１、２１０－２）、ストレージアダプタ（SA）２３０、管理アダプタ（MA）２２０は、それぞれがディスク制御装置２００の筐体に対して脱着可能に独立したユニットとして構成されていることもあるし、これらのうちの２つ以上のアダプタが組み合わされて一つのユニットとして一体化されていることもある。

【００１９】

次に、ディスク制御装置２００の構成要素について詳述する。

図２にホストアダプタ（HA）２１０（２１０－１、２１０－２）のハードウエア構成を示している。ホストアダプタ（HA）２１０（２１０－１、２１０－２）は、サーバ１００（１００－１、１００－２）との間での通信に関する機能を提供する通信インタフェース２１１、ローカルメモリ２１２、フラッシュメモリ等で構成される不揮発性メモリ２１３、ローカルメモリ２１２に記憶されているプログラムを実行することにより当該ホストアダプタ（HA）２１０（２１０－１、２１０－２）の各種機能を実現するマイクロプロセッサ２１４、ホストアダプタ（HA）２１０（２１０－１、２１０－２）とストレージアダプタ（SA）２３０またはキャッシュメモリ（不図示）との間で行われる高速なデータ転送を実現するＩ／Ｏプロセッサ２１５、等を備えている。これらは互いにバス２１６を介して接続されている。不揮発性メモリ２１３には、ホストアダプタ（HA）２１０（２１０－１、２１０－２）が提供する各種の機能を実現するソフトウエアであるマイクロプログラムが記憶されている。マイクロプログラムは適宜ローカルメモリ２１２にロードされてマイクロプロセッサ２１４により実行される。Ｉ／Ｏプロセッサ２１５としては、例えば、ＤＭＡ（Direct Memory Access）プロセッサが用いられる。

【００２０】

図３にストレージアダプタ（SA）２３０の構成を示している。ストレージアダプタ（SA）２３０は、ホストアダプタ（HA）２１０（２１０－１、２１０－２）との間での高速なデータ転送を実現するＩ／Ｏプロセッサ２３１、ローカルメモリ２３２、フラッシュメモリ等で構成される不揮発性メモリ２３３、ローカルメモリ２３２に記憶されているプログラムを実行することにより当該ストレージアダプタ（SA）２３０の各種機能を実現するマイクロプロセッサ２３４、ディスクドライブ２４０（２４０－１～２４０－５）に対するデータの書き込みや読み出しを行うディスクコントローラ２３５等を備える。これらは互いにバス２３６を介して接続されている。不揮発性メモリ２３３には、当該ストレージア

ダプタ（SA）２３０が提供する各種の機能を実現するソフトウエアであるマイクロプログラムが記憶されている。マイクロプログラムは、適宜ローカルメモリ２３２にロードされてマイクロプロセッサ２３４により実行される。Ｉ／Ｏプロセッサ２３１としては、例えば、ＤＭＡ（Direct Memory Access）プロセッサが用いられる。

【００２１】

ストレージアダプタ（SA）２３０は、ホストアダプタ（HA）２１０（２１０－１、２１０－２）が受信した、データの読み出し／書き込み要求を処理する。ディスクドライブ２４０（２４０－１～２４０－５）は、ストレージアダプタ（SA）２３０に接続されている。ストレージアダプタ（SA）２３０は、ディスクドライブ２４０（２４０－１～２４０－５）に対するデータの読み出し／書き込みを行う。ディスクドライブ２４０（２４０－１～２４０－５）は、後述する複製ボリューム（ＬＶ０～ＬＶ２）を構成する物理ボリューム（ＰＤ０～ＰＤ４）を提供する。ディスクコントローラ２３５は、ディスクドライブ２４０（２４０－１～２４０－５）をＲＡＩＤの方式（例えば、ＲＡＩＤ０，１，５）で制御する機能も提供する。

【００２２】

図４に管理アダプタ（MA）２２０のハードウエア構成を示している。管理アダプタ（MA））２２０は、マイクロプロセッサ２２１、メモリ２２２等を備えて構成される。管理アダプタ（MA））２２０は、内部通信インタフェース２２３により内部ネットワーク２５０を介してホストアダプタ（HA）２１０（２１０－１、２１０－２）やストレージアダプタ（SA）２３０と通信可能に接続されている。なお、マイクロプロセッサ２２１、メモリ２２２はそれぞれバス２３５を介して接続されている。管理アダプタ（MA）２２０は、ホストアダプタ（HA）２１０（２１０－１、２１０－２）やストレージアダプタ（SA）２３０に対する各種の設定やディスク制御装置２００における各種障害の監視等を行う。管理アダプタ（MA））２２０は、各ホストアダプタ（HA）２１０（２１０－１、２１０－２）や各ストレージアダプタ（SA）２３０の処理負荷に関する情報を収集することができる。処理負荷に関する情報とは、例えば、マイクロプロセッサ２２１の使用率、各論理ボリュームに対するアクセス頻度等である。これらの情報はホストアダプタ（HA）２１０（２１０－１、２１０－２）やストレージアダプタ（SA）２３０において実行されるプログラムの機能により収集され管理される。管理アダプタ（MA）２２０は、ホストアダプタ（HA）２１０（２１０－１、２１０－２）が受信した設定コマンドに応じた処理を行う。また管理アダプタ（MA）２２０は、管理サーバ１１０に送信する通知コマンドを、内部ネットワーク２５０を介してホストアダプタ（HA）２１０（２１０－１、２１０－２）に引き渡す。

【００２３】

図５は管理サーバ１１０のハードウエア構成を示している。管理サーバ１１０は、ＣＰＵ１１１、メモリ１１２、ポート１１３、記録媒体読取装置１１４、入力装置１１５、出力装置１１６、記憶装置１１７を備える。

ＣＰＵ１１１は管理サーバ１１０の全体の制御を司る。ＣＰＵ１１１は、メモリ１１２に格納されているプログラムを実行することにより当該管理サーバ１１０によって提供される各種の機能を実現する。記録媒体読取装置１１４は、記録媒体１１８に記録されているプログラムやデータを読み取るための装置である。読み取られたプログラムやデータはメモリ１１２や記憶装置１１７に格納される。従って、例えば記録媒体１１８に記録されたプログラムを、記録媒体読取装置１１４を用いて上記記録媒体１１８から読み取って、メモリ１１２や記憶装置１１７に格納するようにすることができる。記録媒体１１８としてはフレキシブルディスクやＣＤ－ＲＯＭ、ＤＶＤ－ＲＯＭ、ＤＶＤ－ＲＡＭ、半導体メモリ等を用いることができる。記憶装置１１７は、例えばハードディスク装置やフレキシブルディスク装置、半導体記憶装置等である。入力装置１１５はオペレータ等により管理サーバ１１０へのデータ入力等のために用いられる。入力装置１１５としては、例えば、キーボードやマウス等が用いられる。出力装置１１６は情報を外部に出力するための装置である。出力装置１１６としては、例えば、ディスプレイやプリンタ等が用いられる。ポート１１３は、例えば、ディスク制御装置２００との通信に用いられ、管理サーバ１１０

はポート１１３を介してディスク制御装置２００のホストアダプタ（HA）２１０（２１０－１、２１０－２）やストレージアダプタ（SA）２３０等と通信を行うことができる。

【００２４】

ストレージシステムの管理者等は、管理サーバ１１０を操作することにより、例えば、ディスクドライブ２４０（２４０－１～２４０－５）に関する各種の設定等を行うことができる。ディスクドライブ２４０（２４０－１～２４０－５）に関する各種の設定としては、例えば、ディスクドライブの増設や減設、ＲＡＩＤ構成の変更（例えばＲＡＩＤ１からＲＡＩＤ５への変更等）等がある。

【００２５】

管理サーバ１１０からは、ストレージシステムの動作状態の確認や故障部位の特定等の作業を行うこともできる。管理サーバ１１０は、ＬＡＮや電話回線等で外部保守センタと接続されている。管理サーバ１１０を利用してストレージシステムの障害監視を行ったり、障害が発生した場合に迅速に対応することが可能である。障害の発生は例えばサーバ１００（１００－１、１００－２）や当該管理サーバ１１０で動作しているオペレーティングシステムやアプリケーション、ドライバソフトウエアなどから通知される。前記通知はＨＴＴＰプロトコルやＳＮＭＰ（Simple Network Management Protocol）、電子メール等により行われる。管理サーバ１１０に対する各種の設定や制御は、管理サーバ１１０で動作するＷｅｂサーバが提供するＷｅｂページを利用して行うこともできる。

【００２６】

次に、ストレージシステムのソフトウエア構成について説明する。図６は本実施例のストレージシステムのソフトウエア構成を示している。この図に示す各部の機能は、各部に対応するハードウエアもしくは各ハードウエアにおいて実行されるプログラムによって実現されている。また図６に示す各種テーブルは、各部に対応するハードウエアもしくは各ハードウエアにおいて実行されるプログラムによって記憶され管理されている。

【００２７】

＝＝＝複製管理機能＝＝＝

まずストレージアダプタ（SA）２３０が備える複製管理機能について説明する。

複製管理機能は、ストレージアダプタ（SA）２３０のマイクロプロセッサ２３４が、不揮発性メモリ２３３に記憶されている複製管理機能を実現するためのプログラムを実行することにより実現される。

本実施例では、図６に示す複製ボリューム管理部６３０が複製管理機能を提供する。複製管理機能は、ある論理ボリューム（以下、「複製元論理ボリューム」と記す）に対するデータの書き込みがあった場合に、その複製元論理ボリュームとは別の論理ボリューム（以下、「複製先論理ボリューム」と記す）にもそのデータを書き込むことにより、ある論理ボリュームに記憶されるデータの複製を他の論理ボリュームにおいても記憶する。一般的なストレージシステムの運用形態では、複製元論理ボリュームは、本番の業務に直接使用されるボリューム（本番ボリューム）として設定され、複製先論理ボリュームは本番ボリュームの複製を管理するためのボリューム（複製ボリューム）に設定される。なお、本実施例においてもそのような設定がなされているものとする。複製元論理ボリュームと複製先論理ボリュームとの対応づけは、上述したようにストレージシステムの管理者等が管理サーバ１１０を操作して設定する。

図７は、複製元論理ボリュームと複製先論理ボリュームとの対応づけが管理される、複製元－複製先管理テーブル７００の一例である。複製元－複製先管理テーブル７００には、複製元論理ボリュームの論理ボリュームＩＤ（ＬＵＮ（Logical Unit Number））に対応させて、複製先論理ボリュームのＬＵＮが対応づけられている。

【００２８】

複製管理機能では、複製元論理ボリュームに対するデータの書き込みがあった場合に複製先論理ボリュームに対してもデータを書き込みが行われるように制御がなされる。上記制御方式には、同期方式と非同期方式とが用意されていることもある。同期方式では複製元論理ボリュームに対するデータの書き込みがあった場合に、複製元論理ボリュームと複

製先論理ボリュームの双方のボリュームにデータが書き込まれた後に情報処理装置に対して書き込完了報告がなされる。つまり、同期方式では、複製元論理ボリュームと複製先論理ボリュームの双方に対する書き込みが完了するまで、情報処理装置には完了報告がなされない。従って、同期方式では複製元論理ボリュームと複製先論理ボリュームの内容の同一性は高い信頼性をもって確保されるが、その分、情報処理装置へのレスポンスの迅速性が損なわれる。一方、非同期方式の場合では、複製元論理ボリュームに対するデータの書き込みがあった場合に、複製先論理ボリュームに対する書き込みが行われたかどうかとは無関係に情報処理装置に完了報告がなされる。従って、非同期方式では情報処理装置へのレスポンスは迅速であるが、その反面、複製元論理ボリュームと複製先論理ボリュームの同一性は必ずしも確保されない。

【００２９】

複製管理機能において、複製元論理ボリュームと複製先論理ボリュームとのペアの関係は「ペア状態」及び「スプリット状態」の２つの状態に適宜移行させることができる。「ペア状態」では、複製元論理ボリュームと複製先論理ボリュームとの間のデータの一致性がリアルタイムに確保されるように制御がなされる。すなわち、複製元論理ボリュームにデータが書き込まれると、上述の同期もしくは非同期方式により複製先論理ボリュームにもデータが書き込まれる。一方、「スプリット状態」は、上記リアルタイムに一致性を確保する制御が解除された状態である。「ペア状態」から「スプリット状態」に移行させることを「スプリット（split）」と称する。逆に「スプリット状態」から「ペア状態」に移行させることを「リシンク（resync）」と称する。

【００３０】

「ペア状態」から「スプリット状態」への移行は、例えば、本番ボリュームのデータのバックアップを取得したり、本番系のデータを開発用やテスト用に使用するといった副業務の目的のために行われる。例えばデータのバックアップを取得しようとする場合には、まず「ペア状態」から「スプリット状態」に移行させてから複製先論理ボリュームのデータをカートリッジテープ等の記録メディアにバックアップする。また開発用やテスト用に本番系のデータを使用したい場合には「ペア状態」から「スプリット状態」に移行させてから複製先論理ボリュームのデータを開発やテスト等に使用する。このように「スプリット状態」に移行させた状態でバックアップ等の副業務がなされることで、本番系業務以外の副業務により本番系業務が影響されるのを極力を抑えることができる。

【００３１】

なお、副業務の完了後などにおいて、「スプリット状態」にあるペアを、再び「ペア状態」に「リシンク」させる場合には、「スプリット」時以降に複製元論理ボリュームに対して行われた更新内容を、複製先論理ボリュームに反映させなければならない。この間の更新差分は、例えばブロック単位で論理ボリューム（以後、差分ボリュームと称する）に記憶される。ペアを「リシンク」する場合、まず複製先論理ボリュームに差分ボリュームの内容を反映させてから、「ペア状態」へと移行させる。

【００３２】

＝＝＝複製ボリュームグループ＝＝＝

次に、複製ボリュームグループについて説明する。一つの複製ボリュームグループには、一つ以上の複製ボリュームが所属する。複製ボリュームグループには、上記「スプリット」時刻以降における複製ボリュームのデータが記憶され、サーバ１００（１００－１、１００－２）からのデータの読み出し／書き込みが禁止されるように属性が設定されているスペアボリューム、及びある時点以降に複製ボリュームに対して行われた差分のデータが記憶される差分ボリュームが適宜含められる。

【００３３】

図８Ａは複製ボリュームグループに関する設定や操作を行うためのコマンドである複製ボリュームグループ操作コマンドのデータフォーマットである。複製ボリュームグループ操作コマンドは、管理サーバ１１０とストレージアダプタ（SA）２３０との間で授受されるコマンドである。図８Ａにおいて、コマンドＩＤ のフィールド８２０には、コマンド

の種類を示す識別子であるコマンドＩＤ が設定される。またコマンド本体のフィールド８３０には、コマンドの種類よって定まるパラメータ等が設定される。コマンドには、複製ボリュームグループについて初期化を指示するコマンドである複製ボリュームグループ初期化コマンド、データ障害が発生した複製ボリュームを障害発生前のデータ内容に復元するように指示するコマンドであるリストアコマンド、指定したスペアボリュームおよび複製ボリュームの現在の属性を問い合わせるコマンドである問合せコマンド等がある。コマンドＩＤのフィールド８２０には、各コマンドに対応したコマンドＩＤ（０：複製ボリュームグループ初期化、１：リストア、３：問合せ（複製ボリューム属性／スペアボリューム属性））が設定される。

【００３４】

図８Ｂにコマンドが上記複製ボリュームグループ初期化コマンドの場合における複製ボリュームグループ操作コマンドのデータフォーマットを一例として示している。図８Ｂにおいて、グループＩＤ のフィールド８３１には、初期化しようする複製ボリュームグループの識別子であるグループＩＤが設定される。複製ボリューム属性リストのフィールド８３２には、さらに当該複製ボリュームグループに所属する各複製ボリュームのＩＤ（ＬＩＤ）（以後、複製ボリュームＩＤと称する）が設定されるフィールド８３４と、複製ボリュームの属性が設定されるフィールド８３５とが含まれる。属性には複製ボリュームに対するアクセスをデータの読出しのみに制限する属性である"Read-Only（RO）"属性と、データの書き込みを許可する属性である"Read-Write（RW）"属性とがある。複製ボリュームが"Read-Only（RO）"属性であれば、フィールド８３４には「RO」が、複製ボリュームが"Read-Write（RW）"属性であれば、フィールド８３４には「RW」がそれぞれ設定される。例えば、バックアップやアーカイブ、ＯＬＡＰ（Online Analytical Processing）等の業務のように、参照目的に使用される複製ボリュームの属性は「RO」に設定される。これに対し、開発やテスト等の書き込みが行われる可能性のある状況で使用される複製ボリュームの属性は「RW」に設定される。一つの複製ボリュームグループ初期化コマンドには、初期化の対象となる複製ボリュームグループに所属する各複製ボリュームに対応する数だけの複製ボリューム属性リスト８３２が含まれる。図８Ｂにおいて、スペア数のフィールド８３３には、当該複製ボリュームグループについて設定されるスペアボリュームの数が設定される。

【００３５】

図８Ｃは、コマンドがリストアコマンドである場合における複製ボリュームグループ操作コマンドのデータフォーマットを示している。リストアコマンド８１０－２には、リストア（復元）の対象となる複製ボリュームの複製ボリュームＩＤ（ＬＩＤと称する）が設定されるフィールド８３６と、リストアしようとするブロック範囲（ＢＩＤｓと称する）が設定されるフィールド８３７とが含まれる。

【００３６】

＝＝＝複製ボリュームグループの初期化＝＝＝

次に管理サーバ１１０からストレージアダプタ（SA）２３０に送信される上述の複製ボリュームグループ初期化コマンドに応じて行われる、複製ボリュームグループを初期化する処理について説明する。以下の説明では、同じ本番ボリュームの同じデータ内容の複製ボリューム（ＬＩＤ＝ＬＶ０～ＬＶ２）を、グループＩＤがＧ０である複製ボリュームグループとして初期化する場合を例として説明する。図９は管理アダプタ（MA）２２０において管理されている複製ボリュームグループ管理表９００である。また図１０は複製ボリュームグループ初期化の処理を説明するフローチャートである。

【００３７】

図１０において、まず管理サーバ１１０の複製ボリュームグループ設定部６１０が、管理アダプタ（MA）２２０の複製ボリュームグループ管理部６２０に対し、図８Ｂに示した複製ボリュームグループ初期化コマンド８１０－１を送信する（S1010）。複製ボリュームグループ管理部６２０は、複製ボリュームグループ初期化コマンド８１０－１を受信する（S1020）。

【００３８】
管理アダプタ（MA）２２０の複製ボリュームグループ管理部６２０は、受信した複製ボリュームグループ初期化コマンド８１０－１に基づいて、複製ボリュームグループ管理表９００の内容を設定する（S1021）。ここで複製ボリュームグループ初期化コマンド８１０－１の複製ボリューム属性リストの中に、「RW」を属性として指定されている複製ボリューム（RW複製ボリューム）が存在する場合には、複製ボリュームグループ管理部６２０は当該RW複製ボリュームについての更新差分を格納する論理ボリューム（差分ボリューム）ＤＬＶ２を設定する。図９の例では、論理ボリュームＩＤ９０２の最下欄の差分ボリューム（ＬＩＤ＝ＤＬＶ２）について対応する論理領域属性の欄９０３の内容が「RW」に設定される。複製ボリュームグループ管理表９００において、リカバリ論理ボリュームＩＤ（リカバリＬＩＤ）の欄９０５は、その複製ボリュームのリカバリに用いられる論理ボリュームのＩＤが設定される。例えば、属性が「RW」である複製ボリュームのリカバリＬＩＤとしては、その複製ボリュームのリカバリのために用いられる差分ボリュームのＩＤが設定される。

【００３９】
複製ボリュームグループ管理部６２０は、図８Ｂに示す複製ボリュームグループ初期化コマンド８１０－１のスペア数のフィールド８３３に設定されている数に相当する分のスペアボリュームを複製ボリュームグループＧ０に対して設定する。以上のようにして複製ボリュームグループ管理表９００の内容が設定される。

【００４０】
次に複製ボリュームグループ管理部６２０は、内容が設定された上記複製ボリュームグループ管理表９００に基づいて、複製ボリューム（ＬＩＤ＝ＬＶ０～ＬＶ２）、差分ボリューム（ＬＩＤ＝ＤＬＶ２）、スペアボリューム（ＬＩＤ＝Ｓ０）について、それぞれ対応する物理ボリューム領域（ＰＤ０～ＰＤ４）の割り当てを行う（S1022）。

【００４１】
次に、複製ボリュームグループ管理部６２０は、複製ボリュームグループ設定部６１０に対し、複製ボリュームグループ初期化コマンド８１０－１に対する応答（リプライ）を送信し（S1023）、論理ボリュームへの読み出し／書き込みを処理するストレージアダプタ（SA）２３０の複製ボリューム管理部６３０に、複製ボリューム初期化コマンド１１５０－１を送信する（S1024）。

【００４２】
図１１Ａに複製ボリューム操作コマンドのデータフォーマットを示している。複製ボリューム操作コマンドには、コマンドの種類を示すコマンドＩＤ（０：複製ボリューム初期化、１：ＰＩＤ変更、２：問合せ（属性／ＰＩＤ／アクセス頻度）、３：リストア）が設定される。このうち「０：複製ボリューム初期化」は、複製ボリュームを初期化するコマンドである。「１：ＰＩＤ変更」は、指定された複製ボリュームの物理ボリュームを変更するコマンドである。「２：問合せ（属性／ＰＩＤ／アクセス頻度）」は、指定された複製ボリュームやスペアボリュームの属性、物理ボリュームＩＤ、アクセス頻度、を問い合わせるコマンドである。「３：リストア」は、指定された複製ボリューム（ＬＩＤ）の指定されたブロック範囲（ＢＩＤｓ）を、指定されたリカバリ論理ボリュームＲＬＩＤを参照してリストアするコマンドである。フィールド１１７０には、コマンドの種類によって内容が異なるコマンド本体が設定される。

【００４３】
図１１Ｂは、複製ボリューム初期化コマンドのデータフォーマットを示している。複製ボリューム初期化コマンド１１５０－１には、コマンドＩＤが設定されるフィールド１１７１と、複製ボリューム、スペアボリューム、差分ボリュームのボリュームＩＤ、属性、物理ボリュームＩＤの組み合わせであるボリュームリストが設定されるフィールド１１７２とが含まれる。

【００４４】
図１１Ｃはリストアコマンド１１５０－２のデータフォーマットを示している。リスト

アコマンド１１５０－２には、複製ボリュームＩＤが設定されるフィールド１１７８、リストアしたいブロック範囲（ＢＩＤｓ）が設定されるフィールド１１７９、リストアのために参照するリカバリＬＩＤ が設定されるフィールド１１８０が含まれる。

【００４５】

＝＝＝複製ボリュームの初期化＝＝＝

次に管理アダプタ（MA）６２０の複製ボリュームグループ管理部６２０と、ストレージアダプタ（SA）２３０の複製ボリューム管理部６３０との間で行われる、複製ボリュームの初期化に関する処理について説明する。図１２はストレージアダプタ（SA）２３０が管理する複製ボリューム管理表１２００を、図１３はストレージアダプタ（SA）２３０が管理する差分管理表１３００をそれぞれ示している。図１４は複製ボリューム初期化の処理を説明するフローチャートである。

【００４６】

図１４において、まず複製ボリュームグループ管理部６２０が、複製ボリューム管理部６３０に対して複製ボリューム初期化コマンド１１５０－１を送信する（S1024）。複製ボリューム管理部６３０は、複製ボリューム初期化コマンド１１５０－１を受信する（S1400）。

【００４７】

次に複製ボリューム管理部６３０は、複製ボリューム初期化コマンド１１５０－１に含まれるボリュームリスト１１７２に基づいて、論理ボリュームＩＤ、論理ボリューム属性、物理ボリュームＩＤ、リカバリ論理ボリュームＩＤを、それぞれ複製ボリューム管理表１２００に設定する（S1401）。なお、複製ボリューム管理表１２００は、複製ボリュームグループ毎に作成される。

【００４８】

図１２に示したように、複製ボリューム管理表１２００には、論理ボリュームＩＤ１２０１、論理ボリュームの属性１２０２、物理ボリュームＩＤ１２０３、リカバリ論理ボリュームＩＤ１２０４に加えて、各論理ボリュームに対するアクセス頻度１２０５も管理されている。なお、このアクセス頻度は、ストレージアダプタ（SA）２３０によって計測される。

【００４９】

次にストレージアダプタ（SA）２３０の複製ボリューム読み出し/書き込み処理部６４０は、複製ボリュームあるいはスペアボリュームに対する読み出し/書き込みのアクセスを処理する度に複製ボリューム管理表１２００のアクセス頻度の欄１２０５に「１」を加算する。管理アダプタ（MA）２２０の複製ボリュームグループ管理部６２０は、問い合わせの対象となる複製ボリュームの複製ボリュームＩＤが設定された複製ボリューム操作コマンド１１５０（コマンドＩＤ＝２）をストレージアダプタ（SA）２３０の複製ボリューム管理部６３０に送信することにより当該複製ボリュームのアクセス頻度を取得することができる。複製ボリュームグループ管理部６２０は、取得したアクセス頻度に基づいて、障害回復処理のために利用する複製ボリュームを選択する。

【００５０】

次に複製ボリューム管理部６３０は、障害回復の対象となる複製ボリュームの属性が「RW」であるかどうかを判定する（S1402）。ここで複製ボリュームの属性が「RW」であるならば、差分管理表１３００にリカバリボリュームＩＤを登録する（S1003）。差分管理表１３００では、登録されている差分ボリューム毎に更新された部分のブロックＩＤを管理している。次に複製ボリューム管理部６３０は、未処理のボリュームリスト１１７２が存在するかどうかを判定し（S1404）、存在すれば（S1401）に進み、存在しなければ複製ボリュームグループ管理部６２０に対して応答（リプライ）を返す（S1405）。

【００５１】

＝＝＝読み出し処理＝＝＝

図１５は本発明の複製ボリューム読み出し/書き込み処理部６４０が行う処理のうち、読出し処理を説明するフローチャートである。

【００５２】
まず複製ボリューム読み出し/書き込み処理部６４０は、ホストアダプタ（HA）２１０から送信されてくる複製ボリュームへの読出し要求を受信する（S1500）。次に複製ボリューム読み出し/書き込み処理部６４０は、前記読出し要求に設定されている複製ボリューム（例えばＬＩＤ＝ＬＶ０の複製ボリューム）からデータを読み出す（S1501）。ストレージアダプタ（SA）２３０は、ディスクドライブ２４０に障害が生じているかどうかをリアルタイムに監視している。S1502において、ストレージアダプタ（SA）２３０は、ドライブ障害が検出されている場合には（S1502：NO）、複製ボリューム管理表１２００の検出されない複製ボリュームのアクセス頻度に１を加算する（S1503）。S1504では、複製ボリューム読み出し/書き込み処理部６４０はホストアダプタ（HA）２１０に前記読出し要求に対する応答を返す。
【００５３】
複製ボリューム読み出し/書き込み処理部６４０は、S1502において、読出し障害を検出すると、読出し障害を検知したことを複製ボリューム障害処理部６５０に通知する（S1510）。複製ボリューム障害処理部６５０は、前記通知を受信すると、図１６のフローチャートに示す読み出し障害処理を実行する。なお、この処理については後述する。
【００５４】
次に複製ボリューム読み出し/書き込み処理部６４０は、複製ボリューム障害処理部６５０から読出し障害のリカバリ結果を受け取り、リカバリが成功したか否かを判定する（S1511）。リカバリが成功した場合には、複製ボリューム読み出し/書き込み処理部６４０は、障害が発生した複製ボリュームからの読出しを再実行し（S1512）、複製ボリューム管理表１２００の当該複製ボリュームのアクセス頻度に１を加算し（S1503）、ホストアダプタ（HA）２１０に読出し要求に対する応答を返す（S1504）。
【００５５】
S1511において、リカバリが成功したか否かを判定した結果、リカバリが失敗した場合には、複製ボリューム読み出し/書き込み処理部６４０は、読出し要求に対する応答として、読出し失敗をホストアダプタ（HA）２１０に返す（S1504）。
【００５６】
＝＝＝読み出し障害処理＝＝＝
図１６は複製ボリュームグループ管理部６２０と複製ボリューム障害処理部６５０との間で行われる読出し障害に関する処理（読み出し障害処理）を説明するフローチャートである。
【００５７】
複製ボリューム障害処理部６５０は、複製ボリューム読み出し／書き込み処理部６４０から送信されてくる（S1510）読出し障害の通知を受信すると、読出し障害が発生した複製ボリュームを複製ボリュームグループ管理部６２０に送信する（S1610）。複製ボリュームグループ管理部６２０は、読出し障害の通知を複製ボリューム障害処理部６５０から受信すると（S1600）、読出し障害処理（S1700）を実行する（S1601）。そして、読出し障害処理（S1700）を実行したリカバリ結果を複製ボリューム障害処理部６５０に送信する（S1602）。
【００５８】
複製ボリューム障害処理部６５０は、リカバリ結果を受信すると（S1611）、リカバリが成功したか否かを判定する（S1612）。判定の結果、リカバリに失敗していた場合には、複製ボリューム障害処理部６５０は、複製ボリューム読み出し／書き込み処理部６４０にリカバリ失敗を通知する（S1613）。一方、S1612の判定において、リカバリが成功していた場合には、複製ボリューム障害処理部６５０は、複製ボリュームグループ管理部６２０から受信したリカバリ結果に基づいて、障害が発生している複製ボリュームを構成している物理ボリュームを設定し直す（S1620）。さらに複製ボリューム障害処理部６５０は、当該物理ボリュームが他の複製ボリュームの物理ボリュームであるか否かを判定し（S1621）、他の複製ボリュームの物理ボリュームの場合には障害が発生した複製ボリューム

のリカバリＬＩＤとして、この複製ボリュームのＬＩＤを設定する（S1622）。当該物理ボリュームがスペアボリュームの物理領域である場合には、当該スペアボリュームを複製ボリューム管理表１２００から削除する（S1623）。そして、複製ボリューム読み出し／書き込み処理部６４０にリカバリに成功した旨を通知する（S1624）。

【００５９】

ここでこのように本実施例のディスク制御装置２００は、障害の内容がハードウエアに関する障害である場合に、障害が生じている複製ボリュームを構成している物理ボリュームに代えて、他の正常に動作している物理ボリュームにより複製ボリュームを構成し直す。これにより、複製ボリュームＬＵＮを変更することなく、複製ボリュームをハードウエア障害から復旧させることができる。

【００６０】

図１７は、図１６の（S1601）において実行される複製ボリューム読出し障害処理（S1600）を説明するフローチャートである。

【００６１】

まず複製ボリュームグループ管理部６２０は、スペアボリュームの有無を判定する（S1701）。ここでスペアボリュームが存在すれば、複製ボリュームグループ管理部６２０は、障害が発生した複製ボリュームの物理ボリュームのＩＤ（ＰＩＤ）をスペアボリュームの物理ボリュームＩＤに変更し（S1710）、そのスペアボリュームＩＤを複製ボリューム管理表１２００から削除する（S1711）。一方、S1701において、スペアボリュームが存在しない場合には、複製ボリュームグループ管理部６２０は、属性が「RO」である複製ボリュームの有無を判定する（S1702）。ここで属性が「RO」である複製ボリュームが存在しない場合には、複製ボリュームグループ管理部６２０は読み出し障害通知を返す（S1703）。一方、属性が「RO」である複製ボリュームが存在する場合には、複製ボリュームグループ管理部６２０は複製ボリューム障害処理部６５０に当該複製ボリュームのアクセス頻度を問い合わせる（S1704）。そして複製ボリュームグループ管理部６２０は、複製ボリューム管理表１２００においてアクセス頻度（Freq）が最小の複製ボリュームを選択し（S1705）、障害が発生している複製ボリュームの物理ボリュームのＩＤを、前記選択した複製ボリュームの物理ボリュームのＩＤに変更する（S1706）。さらに、複製ボリュームグループ管理部６２０は、障害が発生している複製ボリュームのリカバリのための論理ボリュームとして選択した論理ボリュームのＩＤを複製ボリューム管理表１２００に登録する（S1707）。

【００６２】

＝＝＝書き込み処理＝＝＝

図１８は本発明の複製ボリューム読み出し/書き込み処理部６４０によって実行される書き込み処理を説明するフローチャートである。ストレージアダプタ（SA）２３０の複製ボリューム読み出し/書き込み処理部６４０は、ホストアダプタ（HA）２１０から複製ボリュームへの書き込み要求を受信すると（S2200）、その要求に設定されている複製ボリュームの属性が「D-RW」（分離書き込み）であるか否かを判定する（S2201）。判定の結果、属性が「D-RW」であった場合には、複製ボリューム読み出し/書き込み処理部６４０は、図１９に示す分離書き込み処理（S1900）を実行する（S2202）。なお、分離書き込み処理（S1900）については後述する。

【００６３】

一方、前記判定の結果、前記複製ボリュームの属性が「D-RW」でなかった場合（属性が「RW」である場合）には、複製ボリューム読み出し/書き込み処理部６４０は、通常の書き込みを実行する（S2210）。そして、複製ボリューム読み出し/書き込み処理部６４０は、次に通常の書き込みを実行して書き込み障害が発生しているか否かを判定し（S2211）、判定の結果書き込み障害が発生しなかった場合には、複製ボリューム管理表１２００のアクセス頻度（Freq）に１を加算し、ホストアダプタ（HA）２１０に書き込み要求に対する応答（書き込みが成功した旨を示す応答）を返す（S2212）。一方、S2211の判定において、書き込み障害が発生していた場合には、その旨を複製ボリューム障害処理部６５０に

通知する（S2220）。次に複製ボリューム読み出し/書き込み処理部６４０は、複製ボリューム障害処理部６５０から書き込み障害リカバリ結果を受け取るとリカバリが成功したか否かを判定する（S2221）。ここでリカバリに成功していた場合には、（S2201）の処理に進む。またリカバリに失敗していた場合には、書き込み要求に対する応答（書き込みに失敗した旨を示す応答）（リプライ）をホストアダプタ（HA）２１０に返す（S2222）。

【００６４】

図１９は分離書き込み処理を説明するフローチャートである。分離書き込み処理において、まず複製ボリューム読み出し/書き込み処理部６４０は、差分管理表１３００の書き込み先の複製ボリュームの論理ボリュームのＩＤ（ＬＩＤ）について、書き込み先ブロックのＩＤの欄１３５２にブロックＩＤが登録されているか否か（すなわち、これまでに更新がされているか否か）を判定する（S1901）。ここで書き込み先ブロックＩＤの欄１３５２にブロックＩＤが一つも登録されていない場合には、複製ボリューム読み出し/書き込み処理部６４０は、差分管理表１３００の書き込み先ブロックのＩＤの欄１３５２を登録する（S1902）。そして、複製ボリューム読み出し/書き込み処理部６４０は、書き込み先ブロックＩＤに対応するブロックに書き込まれているデータを複製ボリュームから読み出し（S1903）、読み出したデータを書き込みデータデータにより更新し、更新後のデータを差分ボリュームに書き込む（S1904）。

【００６５】

図２０は複製ボリュームグループ管理部６２０と複製ボリューム障害処理部６５０との間で行われる書き込み障害処理を説明するフローチャートである。複製ボリューム障害処理部６５０は、複製ボリューム読み出し／書き込み処理部６４０から送信されてくる書き込み障害の通知を受信すると、書き込み障害が発生している旨を複製ボリュームグループ管理部６２０に送信する（S2010）。

【００６６】

複製ボリュームグループ管理部６２０は、書き込み障害の通知を複製ボリューム障害処理部６５０から受信すると（S2000）、図２１に示す書き込み障害処理（S2100）を実行する（S1701）。書き込み障害処理（S2100）については後述する。次に複製ボリュームグループ管理部６２０は、書き込み障害処理（S2100）を実行した結果であるリカバリ結果を、複製ボリューム障害処理部６５０に送信する（S2002）。

【００６７】

複製ボリューム障害処理部６５０は、複製ボリュームグループ管理６２０から書き込み障害リカバリ結果を受信する（S2011）。複製ボリューム障害処理部６５０は、受信したリカバリ結果に基づいて、リカバリが成功したか否かを判定する（S2012）。ここでリカバリに失敗したと判定した場合には、複製ボリューム障害処理部６５０は、複製ボリューム読み出し／書き込み処理部６４０にリカバリ失敗を通知する（S2013）。

【００６８】

一方、リカバリに成功したと判定した場合には、複製ボリューム障害処理部６５０は、複製ボリュームグループ管理部６２０から受信したリカバリ結果に基づいて、障害が発生している複製ボリュームの物理ボリュームの記憶領域を再設定する（S2020）。さらに複製ボリューム障害処理部６５０は、当該物理ボリュームが他の複製ボリュームの物理ボリュームか否かを判定する（S2021）。ここで前記判定により当該物理ボリュームが他の複製ボリュームの物理ボリュームであった場合には、複製ボリューム障害処理部６５０は、属性を「D-RW」に変更し、障害が発生している複製ボリュームのリカバリ用の複製ボリュームのＩＤ（ＲＬＩＤ）としてその複製ボリュームのＬＩＤを複製ボリュームグループ管理表９００に設定する（S2022）。ここで前記判定により当該物理ボリュームがスペアボリュームを構成している物理ボリュームの記憶領域であった場合には、当該スペアボリュームを複製ボリューム管理表１２００から削除する（S2023）。そして複製ボリューム障害処理部６５０は、複製ボリューム読み出し／書き込み処理部６４０にリカバリに成功した旨を通知する（S2024）。

【００６９】

図２１は、上述の複製ボリューム書き込み障害処理（S2100）を説明するフローチャートである。まず複製ボリュームグループ管理部６２０は、スペアボリュームの存在有無を判定する（S2101）。前記判定の結果、スペアボリュームが存在する場合には、複製ボリュームグループ管理部６２０は、障害が発生している複製ボリュームの物理ボリュームＩＤ（ＰＩＤ）をスペアボリュームの物理ボリュームＩＤに変更し（S2110）、そのスペアボリュームのＩＤを複製ボリューム管理表１２００から削除する（S2111）。一方、前記判定の結果（S2101）、スペアボリュームが存在しない場合には、複製ボリュームグループ管理部６２０は、属性が「RO」である複製ボリュームが存在するかどうかを判定する（S2102）。ここで属性が「RO」である複製ボリュームが存在しない場合には、書き込み障害通知を返す（S2103）。一方、属性が「RO」である複製ボリュームが存在する場合には、複製ボリューム障害処理部６５０に当該複製ボリュームのアクセス頻度を問い合わせる（S2104）。そして、複製ボリュームグループ管理部６２０は、複製ボリューム管理表１２００においてアクセス頻度（Freq）が最小の複製ボリュームを選択し（S2105）、障害が発生した複製ボリュームの属性を「D-RW」に、また物理ボリュームのＩＤ（ＰＩＤ）を、前記選択した複製ボリュームを構成している物理ボリュームのＩＤ（ＰＩＤ）に変更する（S2106）。このように、アクセス頻度が最小の複製ボリュームを用いることにより、復元に関する処理が他の複製ボリュームを利用して行われている業務に与える影響を抑えることができ、複製ボリュームについて可用性を確保することができる。

【００７０】

さらに、複製ボリュームグループ管理部６２０は、複製ボリューム管理表１２００において障害が発生している複製ボリュームのリカバリ論理ボリュームＩＤとして選択した論理ボリュームのＩＤを、複製ボリューム管理表１２００のＲＬＩＤの欄１２０４に登録する（S2107）。

【００７１】

＝＝＝リストア＝＝＝

次に、複製ボリュームグループ設定部６１０と複製ボリュームグループ管理部６２０との間で行われる、障害が発生している複製ボリュームのリストアに関する処理について説明する。図２２は複製ボリュームグループ設定部６１０と複製ボリュームグループ管理部６２０間で行われるリストアに関する処理を説明するフローチャートである。

【００７２】

まず管理サーバ１１０の複製ボリュームグループ設定部６１０から、リストアしたい複製ボリュームのＬＩＤ、リストアしたいブロックＩＤ郡、等が設定された図１１Ｃに示す複製ボリュームリストアコマンド８５０－２が、複製ボリュームグループ管理部６２０に送信される（S2200）。

【００７３】

管理アダプタ（MA）２２０の複製ボリュームグループ管理部６２０は、リストアコマンド８１０－２を受信すると（S2210）、複製ボリュームリストアコマンド設定処理（S2300）を実行する（S2211）。複製ボリュームリストアコマンド設定処理（S1900）の詳細については後述する。次に複製ボリュームグループ管理部６２０は、複製ボリュームリストアコマンド８５０－２の設定に成功したか否かを判定し（S2212）、設定に成功したと判定した場合には、複製ボリューム障害処理部６５０にリストアコマンド１１５０－２を送信する（S2213）。また複製ボリュームグループ管理部６２０は、リストア結果を管理サーバ１１０の複製ボリュームグループ設定部６１０に送信する（S2214）。

【００７４】

図２３は上述のリストアコマンド設定処理（S2300）を説明するフローチャートである。なお、この処理は管理アダプタ（MA）２２０の複製ボリュームグループ管理部６２０と、ストレージアダプタ（SA）２３０の複製ボリューム障害処理部６５０との間で行われる。

【００７５】

まず複製ボリュームグループ管理部６２０は、スペアボリュームが存在するかどうかを

判定する（S2301）。ここでスペアボリュームが存在する場合には、複製ボリュームグループ管理部６２０は、リストアコマンド１１５０－２に設定されているリカバリＬＩＤを、複製ボリュームグループ管理表９００の該当のＬＩＤの欄９０２に設定する（S1910）。一方、スペアボリュームが存在しない場合には、複製ボリュームグループ管理部６２０は、属性が「RO」である複製ボリュームが存在するかどうかを判定する（S2302）。ここで属性が「RO」である複製ボリュームが存在しない場合には、リストアに失敗した旨の通知を返す（S2303）。一方、属性が「RO」である複製ボリュームが存在する場合には、複製ボリュームグループ管理部６２０は、複製ボリューム障害処理部６５０に当該複製ボリュームのアクセス頻度を問い合わせる（S2304）。そして、複製ボリュームグループ管理部６２０は、問い合わせの結果として送られてくるアクセス頻度に基づいて、複製ボリューム管理表１２００におけるアクセス頻度（Freq）が最小の複製ボリュームを選択し（S2305）、リストアコマンド１１５０－２に設定されているリカバリＬＩＤを、選択した複製ボリュームとして設定する（S2306）。このように、アクセス頻度が最小の複製ボリュームを用いることにより、復元に関する処理が他の複製ボリュームを利用して行われている業務に与える影響を抑えることができ、複製ボリュームについて可用性を確保することができる。

【００７６】

図２４は本発明の複製ボリュームリストア処理を説明するフローチャートである。

複製ボリューム障害処理部６５０は、複製ボリュームグループ管理部６２０からリストアコマンド１１５０－２を受信すると（S2400）、リストアコマンド１１５０－２に設定されているリストア対象の複製ボリュームの属性が「RW」であるか否かを判定する（S2401）。ここで属性が「RW」でない場合には、複製ボリューム障害処理部６５０は、リストアコマンド１１５０－２に設定されているリカバリボリュームから、リストアコマンド１１５０－２に設定されているブロックのデータを読出し、読み出したデータをリストアしたい複製ボリュームに書き込む（S2410）。この書き込みが完了すると、複製ボリューム障害処理部６５０は、リストアに終了した旨の通知を複製ボリュームグループ管理部６２０に送信する（S2404）。一方、S2401の判定において、属性が「RW」であった場合には、複製ボリューム障害処理部６５０は、リストアコマンド１１５０－２が指定するリストアしたいブロック範囲が更新されているか否かを、差分管理表１３００を参照することにより判定する（S2402）。前記判定の結果、更新されていないと判定された場合には、S2410に進む。一方、更新されていると判定された場合には、複製ボリューム障害処理部６５０は、リストアしたい複製ボリュームの差分ボリュームからリストアしたいブロックのデータを読出し、読み出したデータをリストアしたい複製ボリュームに書き込む（S2403）。前記書き込みが終了した後、複製ボリューム障害処理部６５０は、リストアが終了した旨を複製ボリュームグループ管理部６２０に送信する（S2404）。

【００７７】

以上実施例を示して説明したように、本発明によれば複製ボリュームの可用性を確保することができる。また他の複製ボリュームや後述するスペアボリュームを用いて複製ボリュームを復元するようにしているため、例えば、データ分析や開発、テスト等において必要とされる必要な時点（ポジション）における複製ボリュームの内容の複製ボリュームを復元することが可能である。

【００７８】

また複製ボリュームの復元は必ずしも一律の方法で行われるわけではなく、障害の内容に応じた方法により行われるので、効率よく柔軟な方法で複製ボリュームを復元することができる。また上記復元に際して使用する読み出し専用ボリュームとしてアクセス頻度の最小のものを用いることにより、復元に関する処理が他の複製ボリュームを利用して行われている業務に与える影響を抑えることができ、複製ボリュームについて可用性を確保することができる。さらに上記復元に際して使用する読み出し専用ボリュームに代えて、ある時刻における前記複製ボリュームの内容に維持され、かつ、前記外部からのアクセス要求に対するデータの読み出し及び書き込みを禁止するように制御される論理ボリューム（

スペアボリューム）を用いることにより、復元に関する処理が他の複製ボリュームを利用して行われている業務に与える影響をより少なくすることができ、複製ボリュームについて高い可用性を確保することができる。さらにハードウエアに関する障害である場合には、障害が生じている複製ボリュームを構成している記憶領域を供給している記憶装置に代えて、他の正常に動作している記憶装置が供給している記憶領域を用いて前記記憶領域を構成し直す

なお、昨今、複製ボリューム等に使用するディスクドライブについて、データの管理コストを低減すべく、ＡＴＡドライブ等の安価なドライブが採用されることがあるが、安価なドライブを利用したことにより障害発生頻度が向上すれば復元作業も増えることになりそのための管理コストが増えることになる。従って安価なドライブを使用してＴＣＯ（Total Cost Of Ownership）削減を実現するには、管理コストの低減が不可欠であるが、本発明によれば、効率よく複製ボリュームを復元することができるため、安価なドライブを用いてＴＣＯ削減を実現することが可能となる。

【００７９】

なお、以上の説明は本発明の理解を容易にするためのものであり、本発明を限定するものではない。本発明はその趣旨を逸脱することなく変更、改良され得ると共に本発明にはその等価物が含まれることは勿論である。

【図面の簡単な説明】

【００８０】

【図１】本発明のストレージシステムの概略的なハードウエア構成を示す図である。

【図２】本発明のホストアダプタ（HA）２１０のハードウエア構成を示す図である。

【図３】本発明のストレージアダプタ（SA）２３０のハードウエア構成を示す図である。

【図４】本発明の管理アダプタ（MA）２２０のハードウエア構成を示す図である。

【図５】本発明の管理サーバ１１０のハードウエア構成を示す図である。

【図６】本実施例のストレージシステムの主なソフトウエア構成を示す図である。

【図７】本発明の一実施例による複製元－複製先管理テーブル７００を示す図である。

【図８】図８Ａは本発明の一実施例による複製ボリュームグループに関する設定や操作を行うためのコマンドである複製ボリュームグループ操作コマンドのデータフォーマットを示す図であり、図８Ｂは本発明の一実施例による複製ボリュームグループ初期化コマンドのデータフォーマットを示す図であり、図８Ｃは本発明の一実施例によるリストアコマンドのデータフォーマットを示す図である。

【図９】本発明の一実施例による複製ボリュームグループ管理表９００を示す図である。

【図１０】本発明の一実施例による複製ボリュームグループ初期化の処理を説明するフローチャートを示す図である。

【図１１】図１１Ａは本発明の一実施例による複製ボリューム操作コマンドのデータフォーマットを示す図であり、図１１Ｂは本発明の一実施例による複製ボリューム初期化コマンド１１５０－１のデータフォーマットを示す図であり、図１１Ｃは本発明の一実施例によるリストアコマンド１１５０－２のデータフォーマットを示す図である。

【図１２】本発明の一実施例によるストレージアダプタ（SA）２３０が管理する複製ボリューム管理表１２００を示す図である。

【図１３】本発明の一実施例による差分管理表１３００を示す図である。

【図１４】本発明の一実施例による複製ボリューム初期化の処理を説明するフローチャートを示す図である。

【図１５】本発明の一実施例による読出し処理を説明するフローチャートを示す図である。

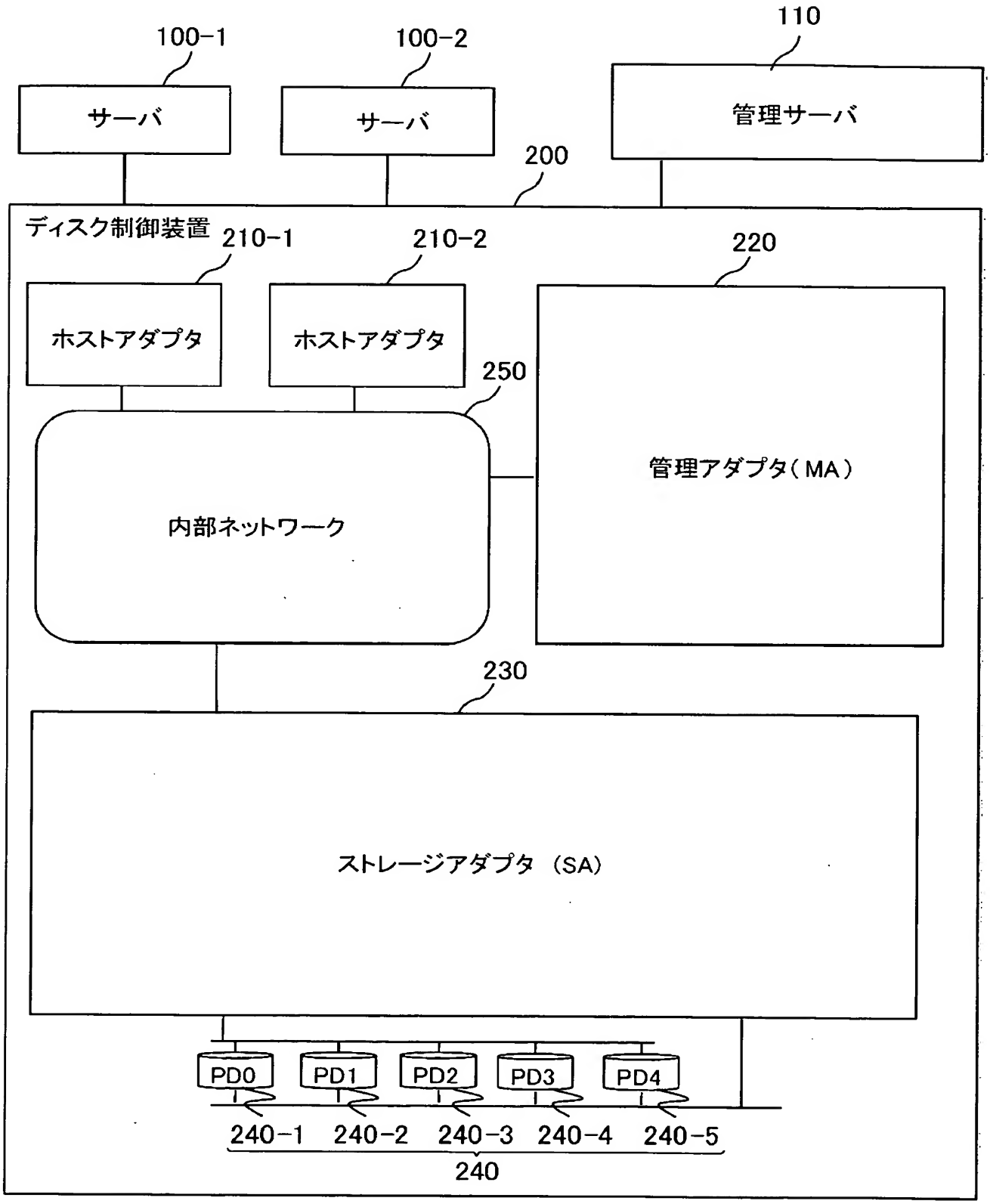
【図１６】本発明の一実施例による読み出し障害処理を説明するフローチャートを示

す図である。
【図１７】本発明の一実施例による複製ボリューム読出し障害処理を説明するフローチャートを示す図である。
【図１８】本発明の一実施例による書き込み処理を説明するフローチャートを示す図である。
【図１９】本発明の一実施例による分離書き込み処理を説明するフローチャートを示す図である。
【図２０】本発明の一実施例による書き込み障害処理を説明するフローチャートを示す図である。
【図２１】本発明の一実施例による複製ボリューム書き込み障害処理を説明するフローチャートを示す図である。
【図２２】本発明の一実施例によるリストアに関する処理を説明するフローチャートを示す図である。
【図２３】本発明の一実施例によるリストアコマンド設定処理を説明するフローチャートを示す図である。
【図２４】本発明の一実施例による複製ボリュームリストア処理を説明するフローチャートを示す図である。

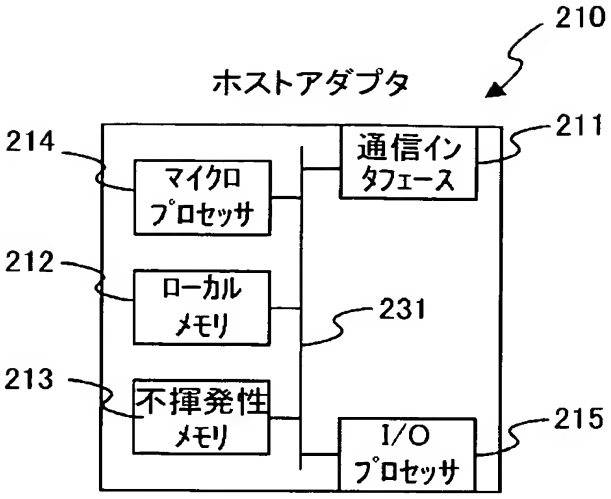
【符号の説明】

【００８１】
１００　サーバ
１１０　管理サーバ
２００　ディスク制御装置
２１０　ホストアダプタ（HA）
２３０　ストレージアダプタ（SA）
２４０　ディスクドライブ（磁気ディスク装置）
２５０　内部ネットワーク
６１０　複製ボリュームグループ設定部
６２０　複製ボリュームグループ管理部
６３０　複製ボリューム管理部
６４０　複製ボリュームＲ／Ｗ処理部
７００　複製元－複製先管理テーブル
９００　複製ボリュームグループ管理表
１２００　複製ボリューム管理表
１３００　差分管理表

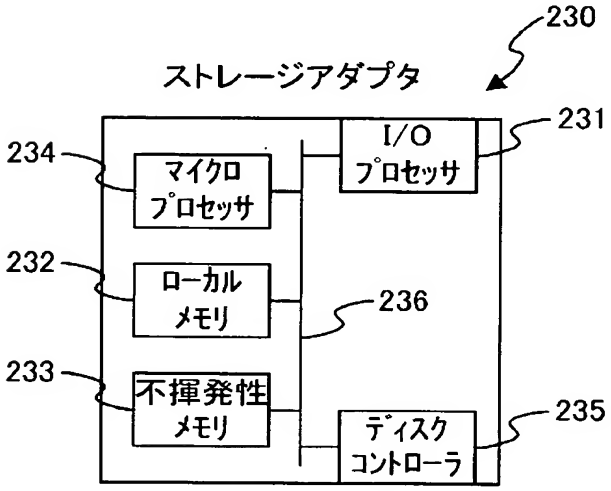
【書類名】図面
【図１】


100-1
100-2
110
サーバ
サーバ
管理サーバ
200
ディスク制御装置
210-1
210-2
220
ホストアダプタ
ホストアダプタ
250
内部ネットワーク
管理アダプタ（MA）
230
ストレージアダプタ （SA）
PD0
PD1
PD2
PD3
PD4
240-1
240-2
240-3
240-4
240-5
240

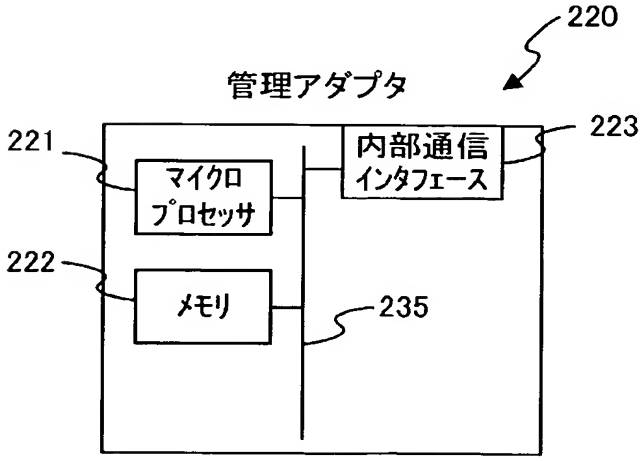
【図２】


210
ホストアダプタ
214
マイクロ
プロセッサ
通信イン
タフェース
211
212
ローカル
メモリ
231
213
不揮発性
メモリ
I/O
プロセッサ
215

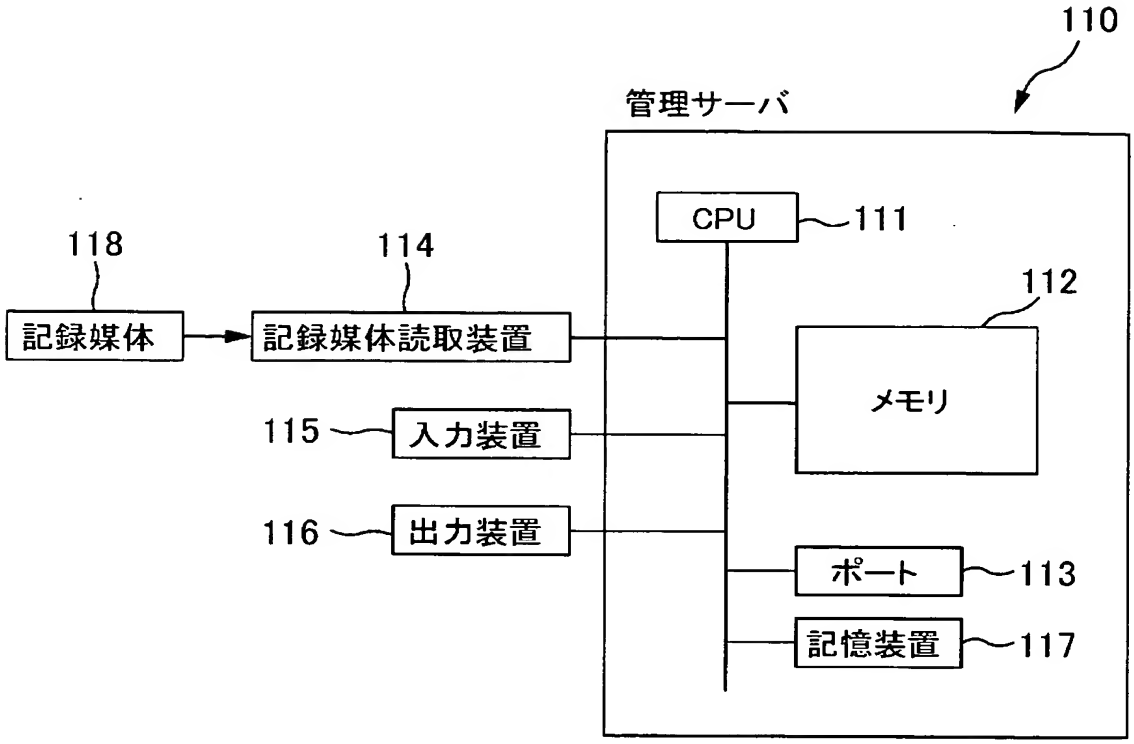
【図３】


230
ストレージアダプタ
234
マイクロ
プロセッサ
I/O
プロセッサ
231
232
ローカル
メモリ
236
233
不揮発性
メモリ
ディスク
コントローラ
235

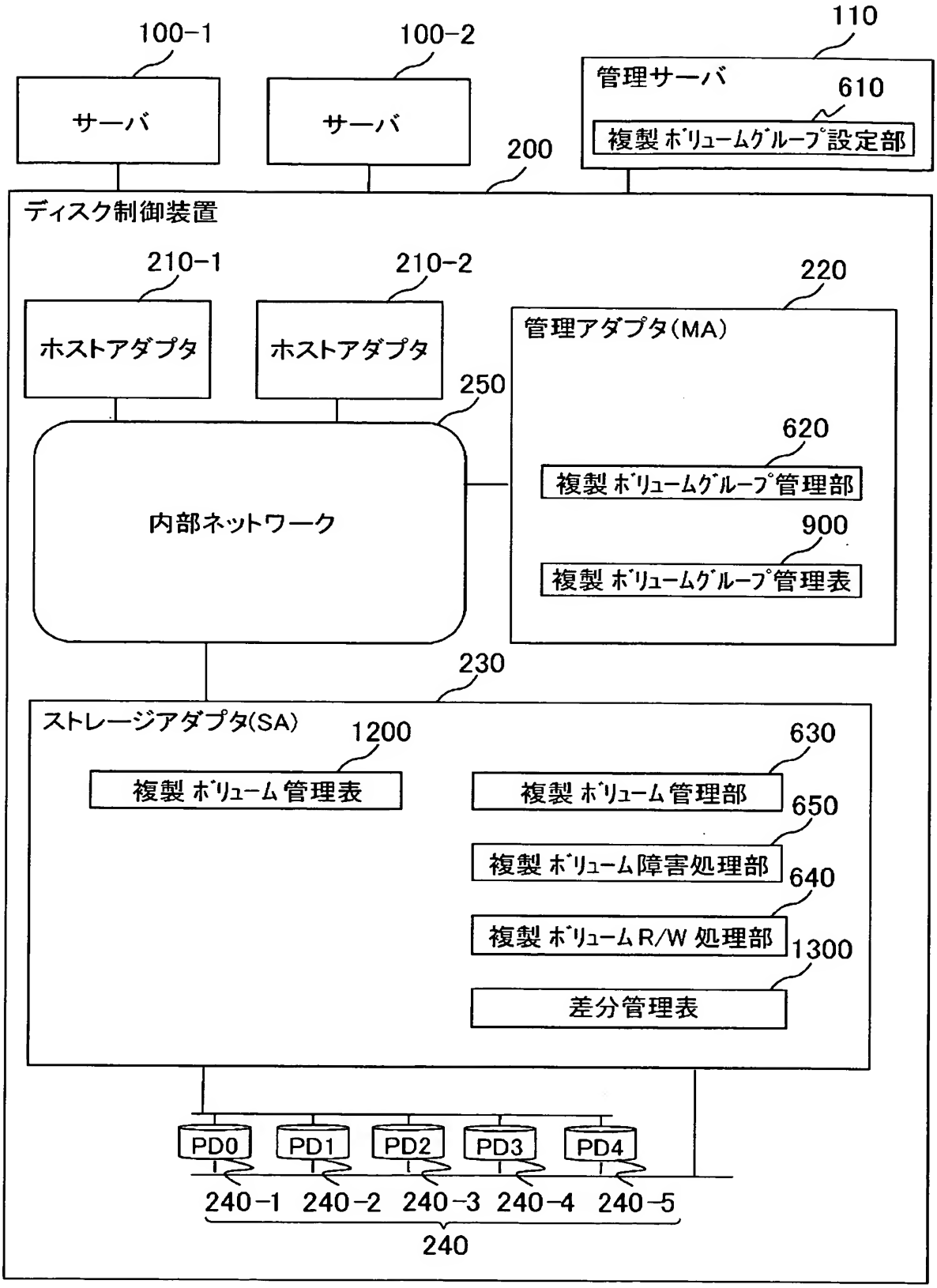
【図４】


220
管理アダプタ
221
マイクロ
プロセッサ
内部通信
インタフェース
223
222
メモリ
235

【図５】


110
管理サーバ
CPU
111
118
114
112
記録媒体
記録媒体読取装置
メモリ
115
入力装置
116
出力装置
ポート
113
記憶装置
117

【図６】


100-1
100-2
110
サーバ
サーバ
管理サーバ
610
複製ボリュームグループ設定部
200
ディスク制御装置
210-1
210-2
220
ホストアダプタ
ホストアダプタ
管理アダプタ(MA)
250
620
内部ネットワーク
複製ボリュームグループ管理部
900
複製ボリュームグループ管理表
230
ストレージアダプタ(SA)
1200
630
複製ボリューム管理表
複製ボリューム管理部
650
複製ボリューム障害処理部
640
複製ボリュームR/W処理部
1300
差分管理表
PD0
PD1
PD2
PD3
PD4
240-1
240-2
240-3
240-4
240-5
240

【図７】

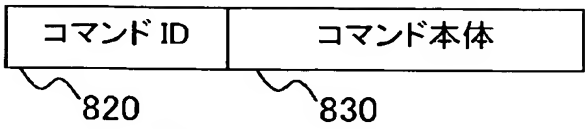
700

複製元-複製先管理テーブル

複製元LU (LUN)	複製先LU (LUN)
1	10
2	11
3	12
4	13
5	14

【図８】

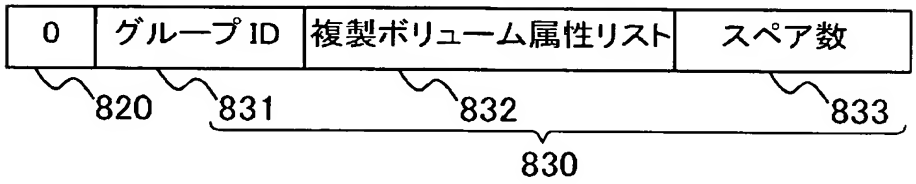
複製ボリュームグループ操作コマンド（管理サーバ/管理アダプタ）



0:複製ボリュームグループ初期化
1:リストア
3:問合せ（複製ボリューム属性/スペアボリューム属性）

図8A

複数ボリュームグループ初期化コマンド 810-1



複数ボリューム属性リスト832

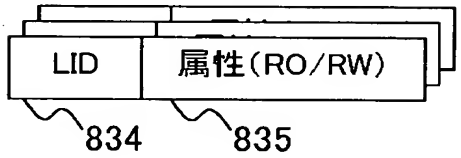


図8B

リストア・コマンド 810-2

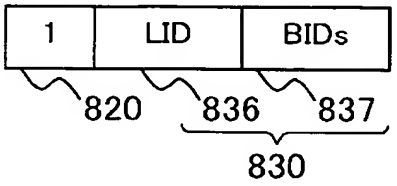
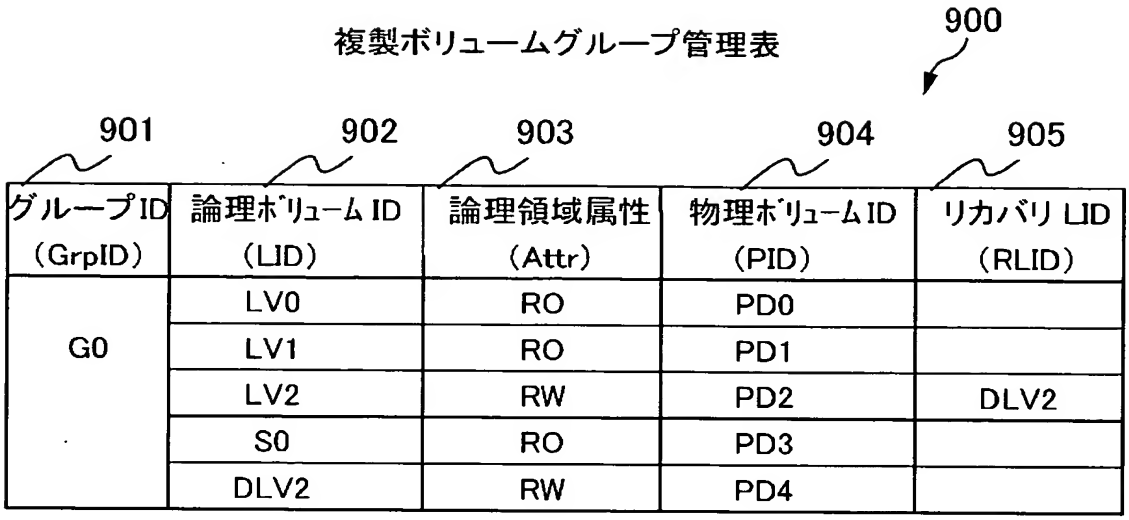


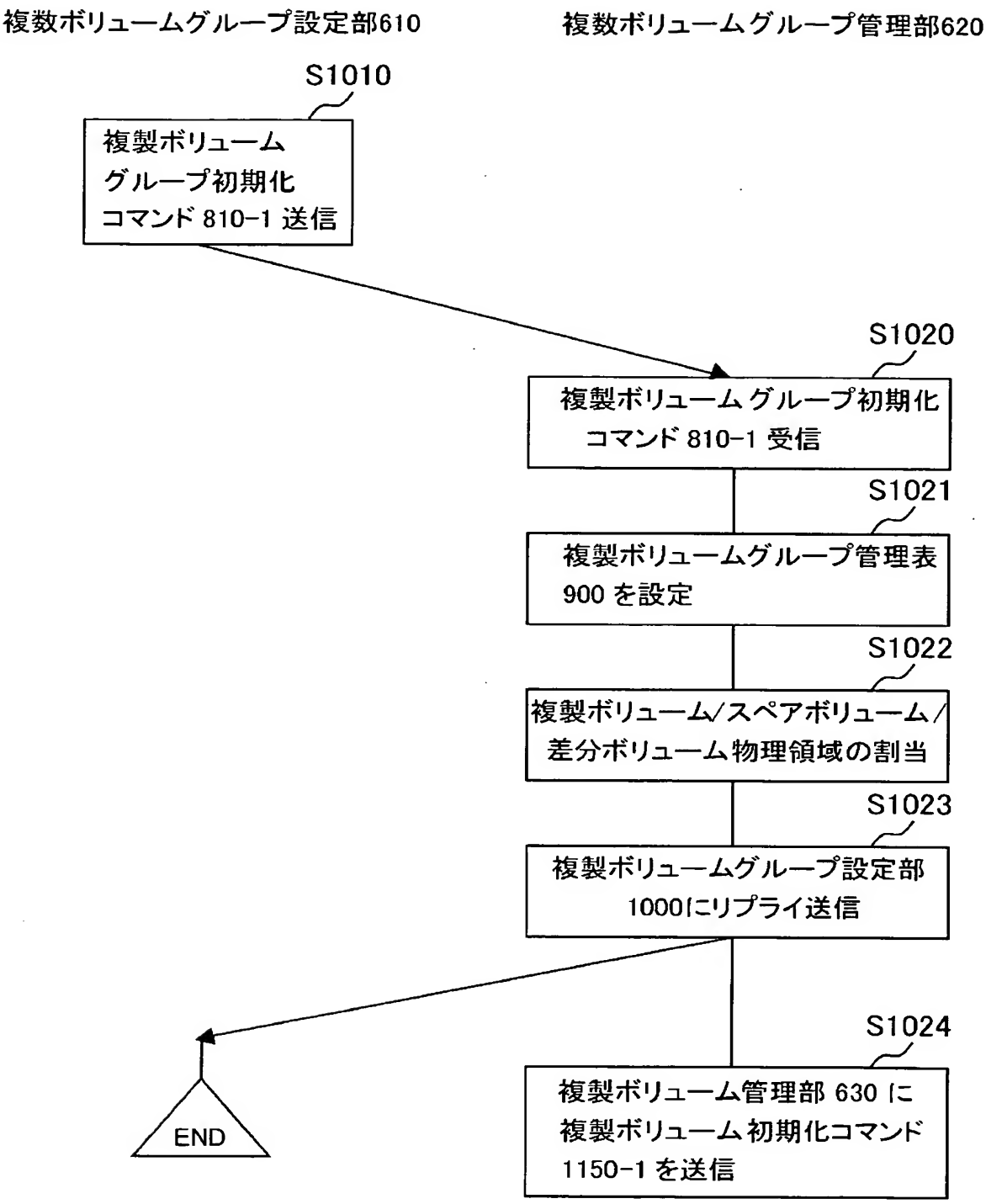
図8C

【図９】



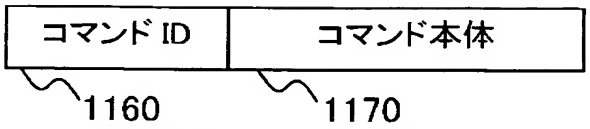
グループID (GrpID)	論理ボリューム ID (LID)	論理領域属性 (Attr)	物理ボリューム ID (PID)	リカバリ LID (RLID)
G0	LV0	RO	PD0	
	LV1	RO	PD1	
	LV2	RW	PD2	DLV2
	S0	RO	PD3	
	DLV2	RW	PD4	

【図１０】


複数ボリュームグループ設定部610
複数ボリュームグループ管理部620
S1010
複製ボリューム
グループ初期化
コマンド 810-1 送信
S1020
複製ボリューム グループ初期化
コマンド 810-1 受信
S1021
複製ボリュームグループ管理表
900 を設定
S1022
複製ボリューム/スペアボリューム/
差分ボリューム物理領域の割当
S1023
複製ボリュームグループ設定部
1000にリプライ送信
S1024
複製ボリューム管理部 630 に
複製ボリューム初期化コマンド
1150-1 を送信
END

【図１１】

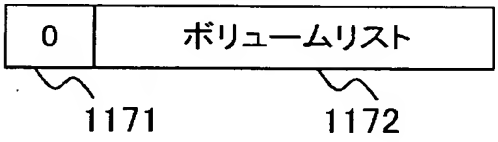
複数ボリューム操作コマンド（ストレージアダプタ/管理アダプタ）1150



0:複製ボリューム 初期化
1:PID 変更
2:問合せ（属性/PID/アクセス頻度）
3:リストア

図11A

複製ボリューム初期化コマンド1150-1



ボリュームリスト1172

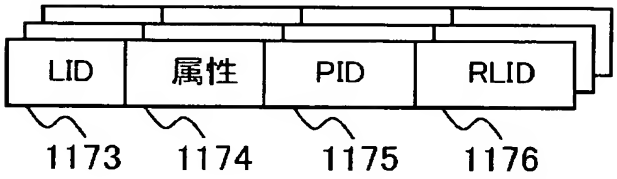


図11B

複製ボリュームリストアコマンド1150-2

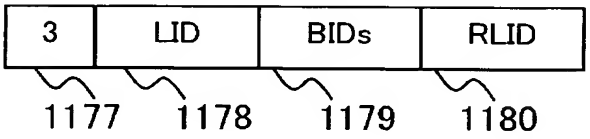


図11C

【図１２】

複製ボリューム 管理表

1200

1201	1202	1203	1204	1205
論理ボリュームID（LID）	論理ボリューム属性（Attr）	物理ボリュームID（PID）	リカバリ LID（RLID）	アクセス頻度（Freq）
LV0	RO	PD0		
LV1	RO	PD1		
LV2	RW	PD2	DLV2	
S0	—	PD3		
DLV2	RW	PD4		

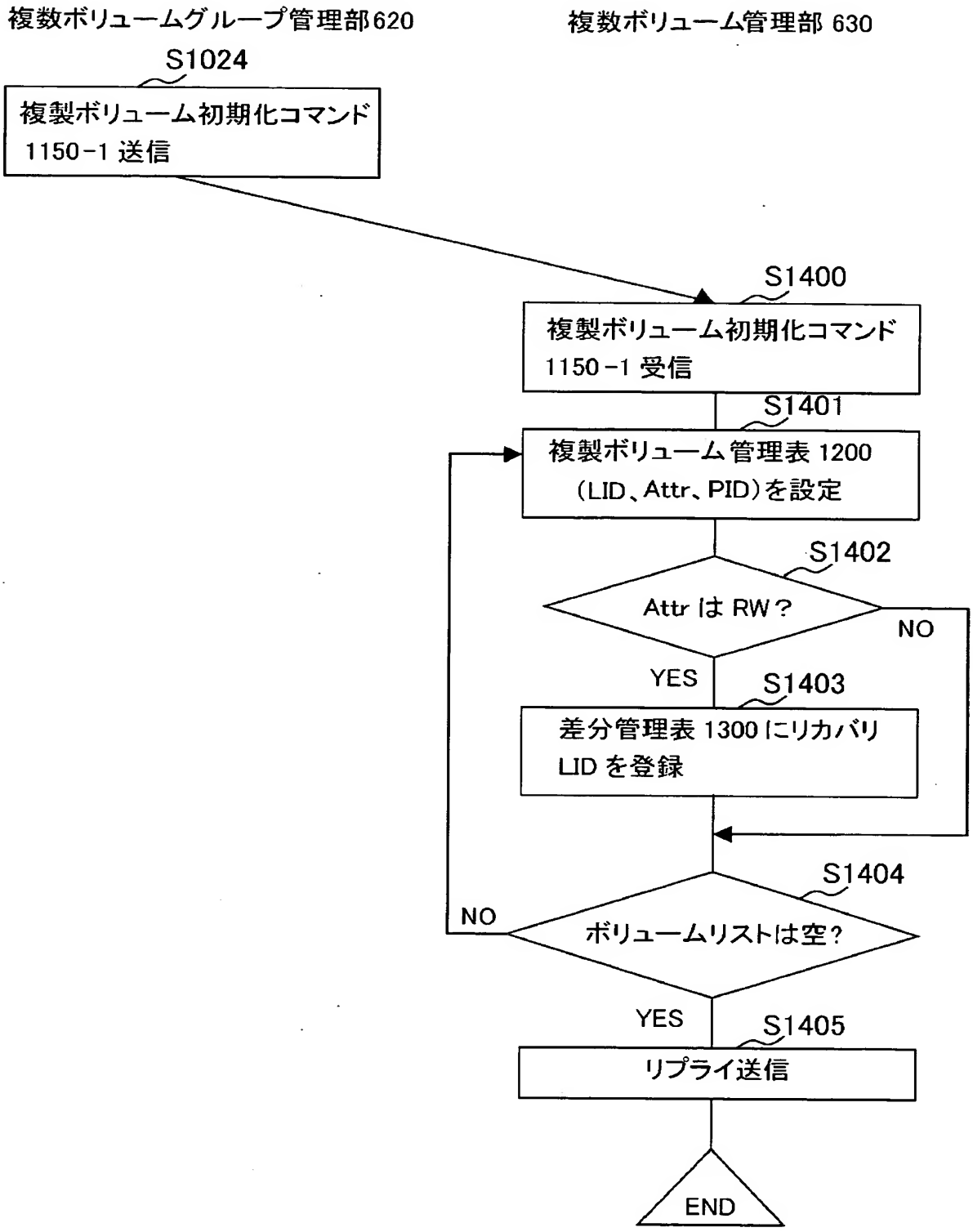
【図１３】

差分管理表

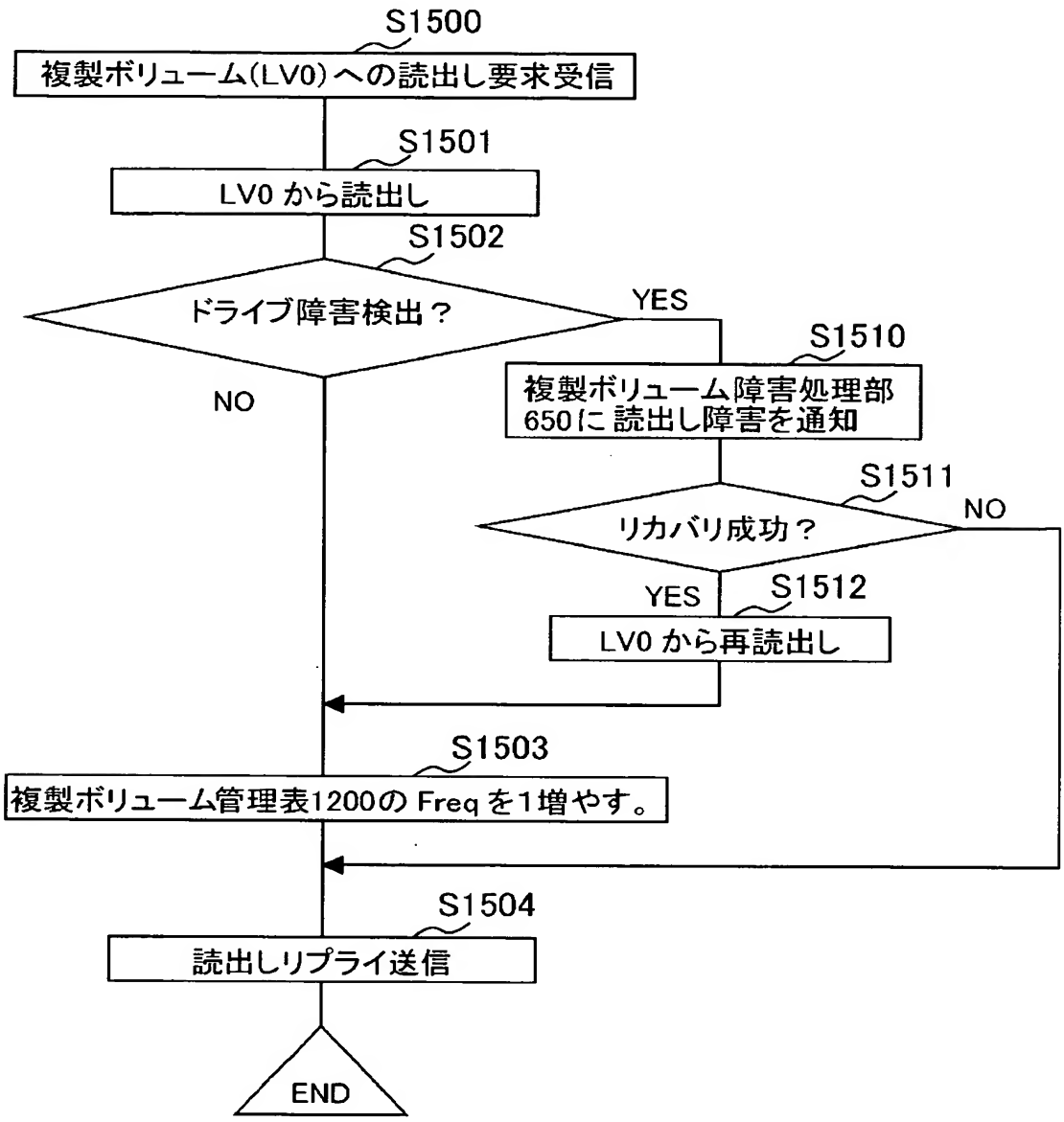
1300

1351	1352
差分 LID（DLID）	ブロック ID（BID）
DLV2	

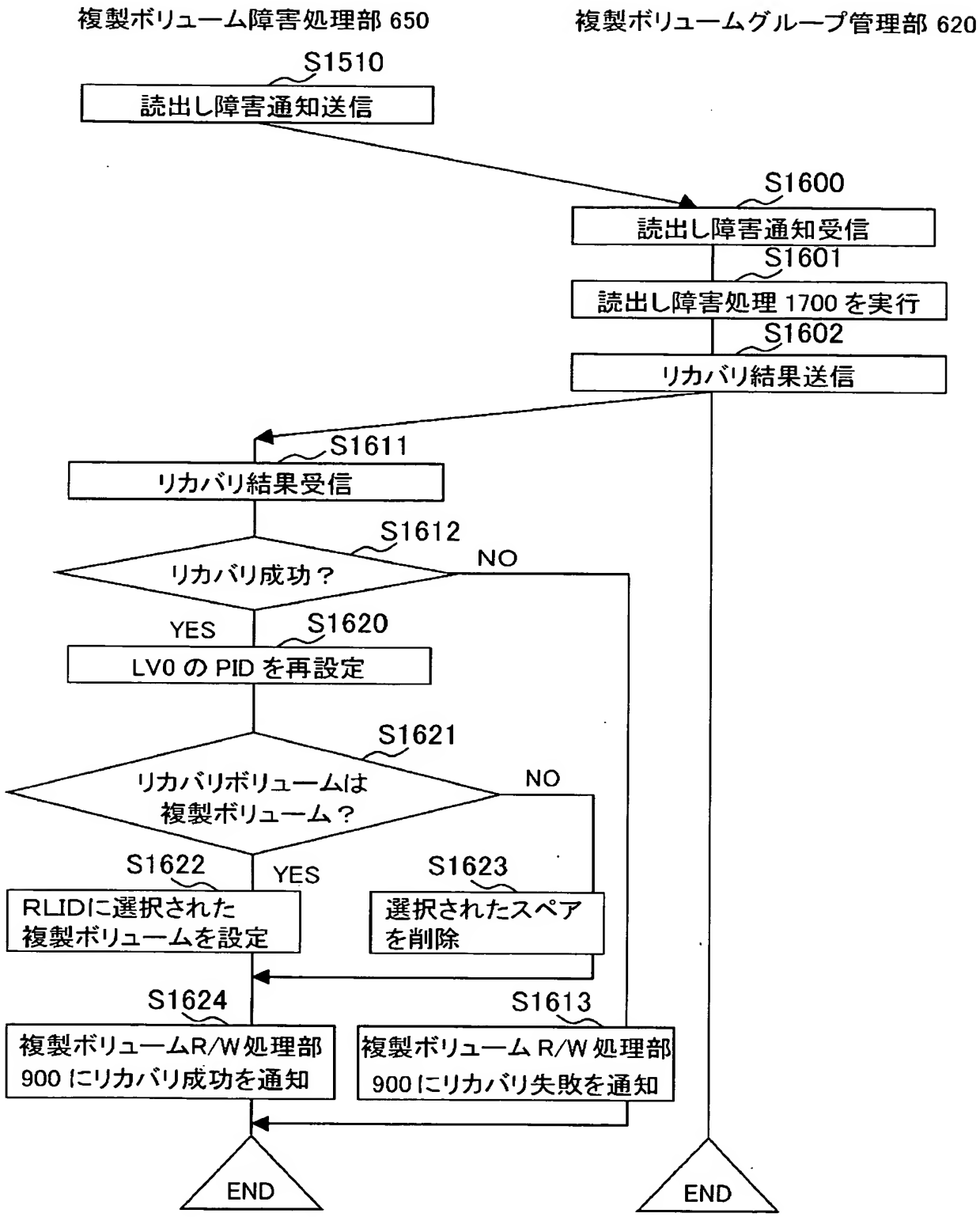
【図１４】


複数ボリュームグループ管理部620
複数ボリューム管理部 630
S1024
複製ボリューム初期化コマンド
1150-1 送信
S1400
複製ボリューム初期化コマンド
1150-1 受信
S1401
複製ボリューム管理表 1200
(LID、Attr、PID)を設定
S1402
Attr は RW？
NO
YES
S1403
差分管理表 1300 にリカバリ
LID を登録
S1404
NO
ボリュームリストは空？
YES
S1405
リプライ送信
END

【図１５】

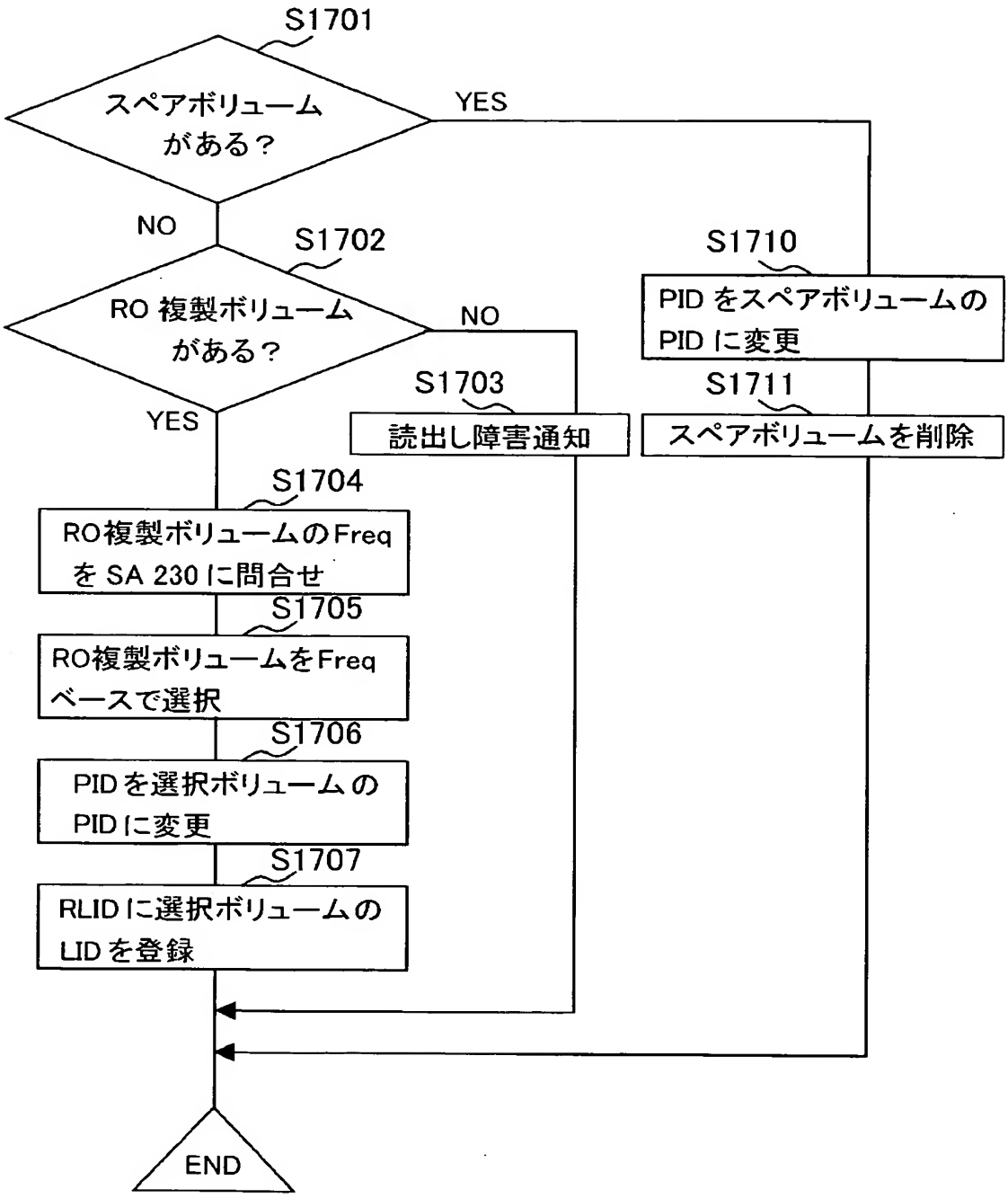

複製ボリューム R/W 処理部　640 (読出し処理)
S1500
複製ボリューム(LV0)への読出し要求受信
S1501
LV0 から読出し
S1502
ドライブ障害検出？
YES
NO
S1510
複製ボリューム障害処理部
650に 読出し障害を通知
S1511
リカバリ成功？
NO
YES
S1512
LV0 から再読出し
S1503
複製ボリューム管理表1200の Freq を1増やす。
S1504
読出しリプライ送信
END

【図１６】

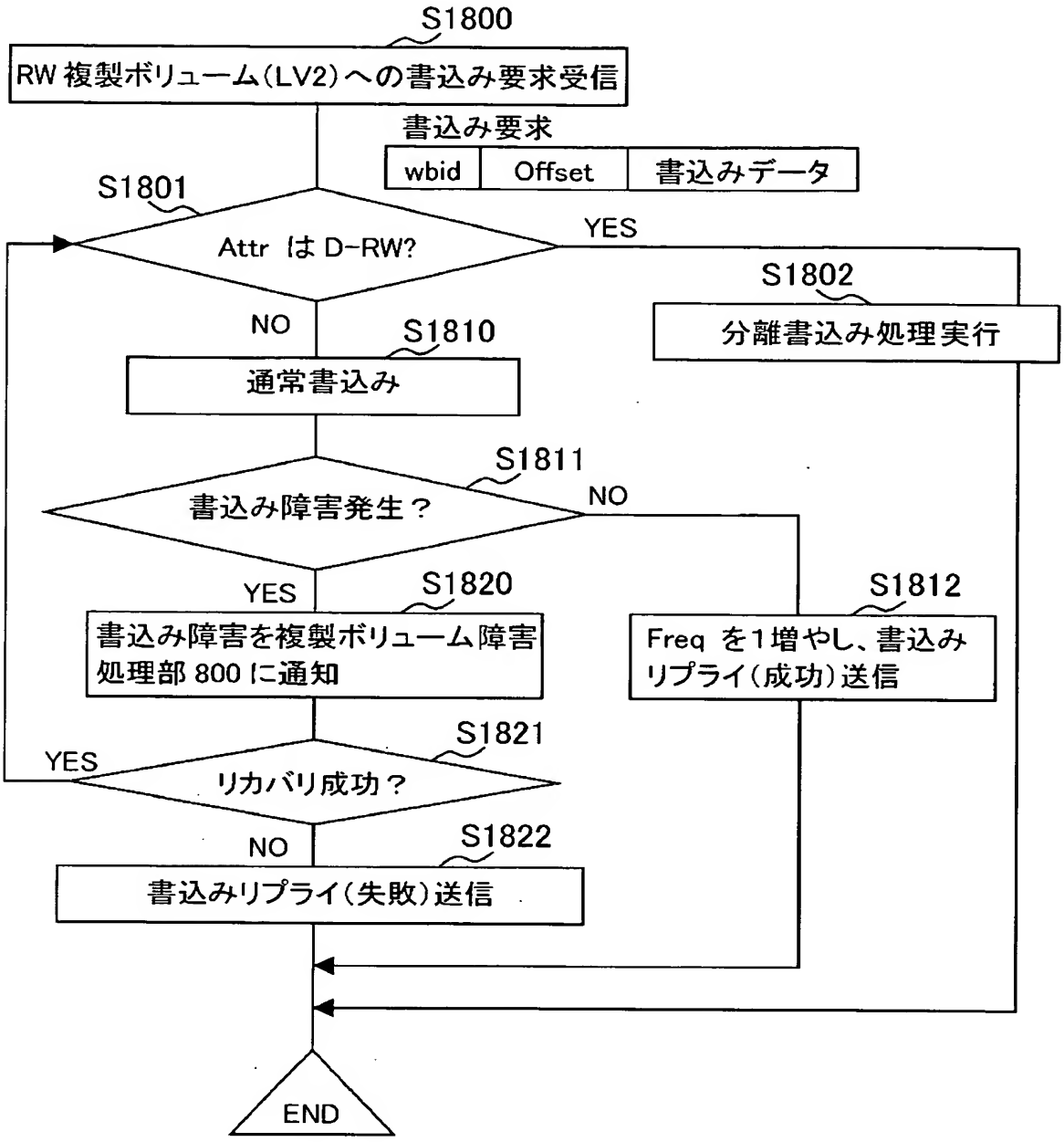

複製ボリューム障害処理部 650
複製ボリュームグループ管理部 620
S1510
読出し障害通知送信
S1600
読出し障害通知受信
S1601
読出し障害処理 1700 を実行
S1602
リカバリ結果送信
S1611
リカバリ結果受信
S1612
リカバリ成功？
NO
YES
S1620
LV0 の PID を再設定
S1621
リカバリボリュームは
複製ボリューム？
NO
S1622
YES
S1623
RLIDに選択された
複製ボリュームを設定
選択されたスペア
を削除
S1624
S1613
複製ボリュームR/W処理部
900 にリカバリ成功を通知
複製ボリューム R/W 処理部
900 にリカバリ失敗を通知
END
END

【図１７】

複製ボリューム読出し障害処理

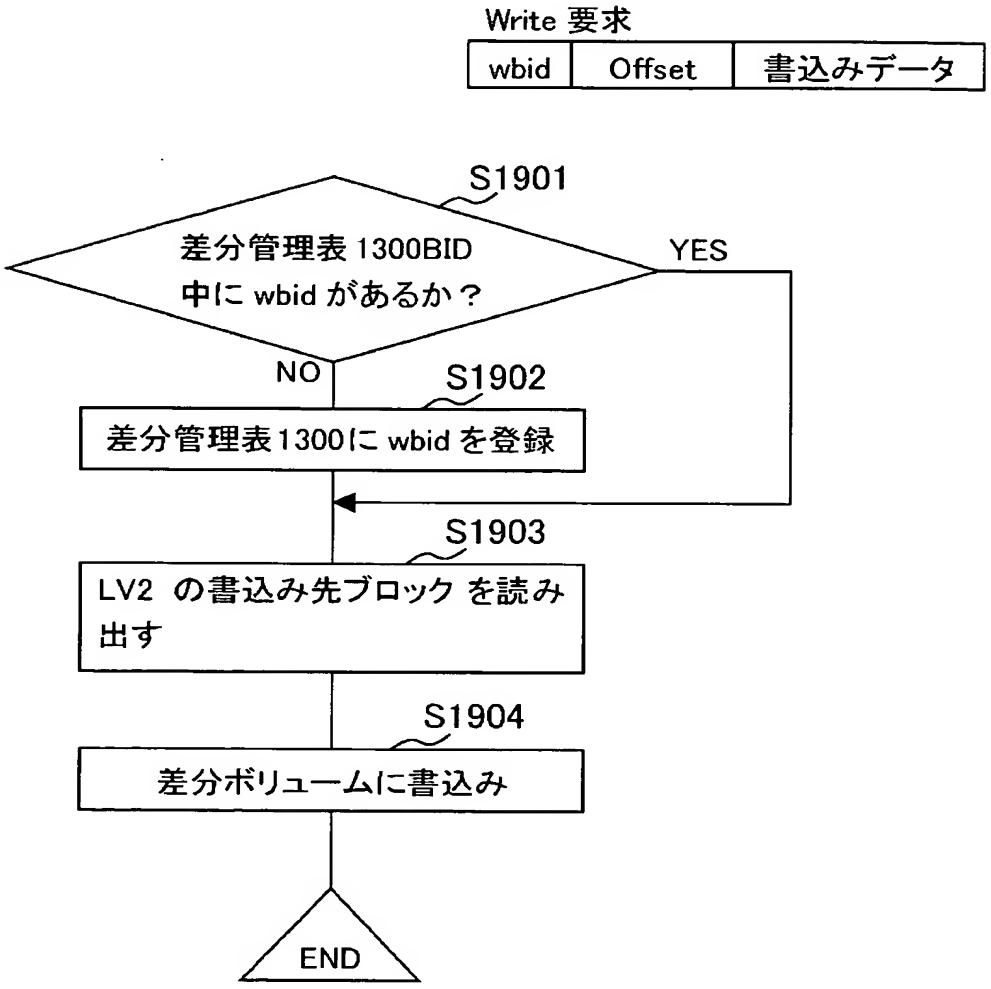


【図１８】

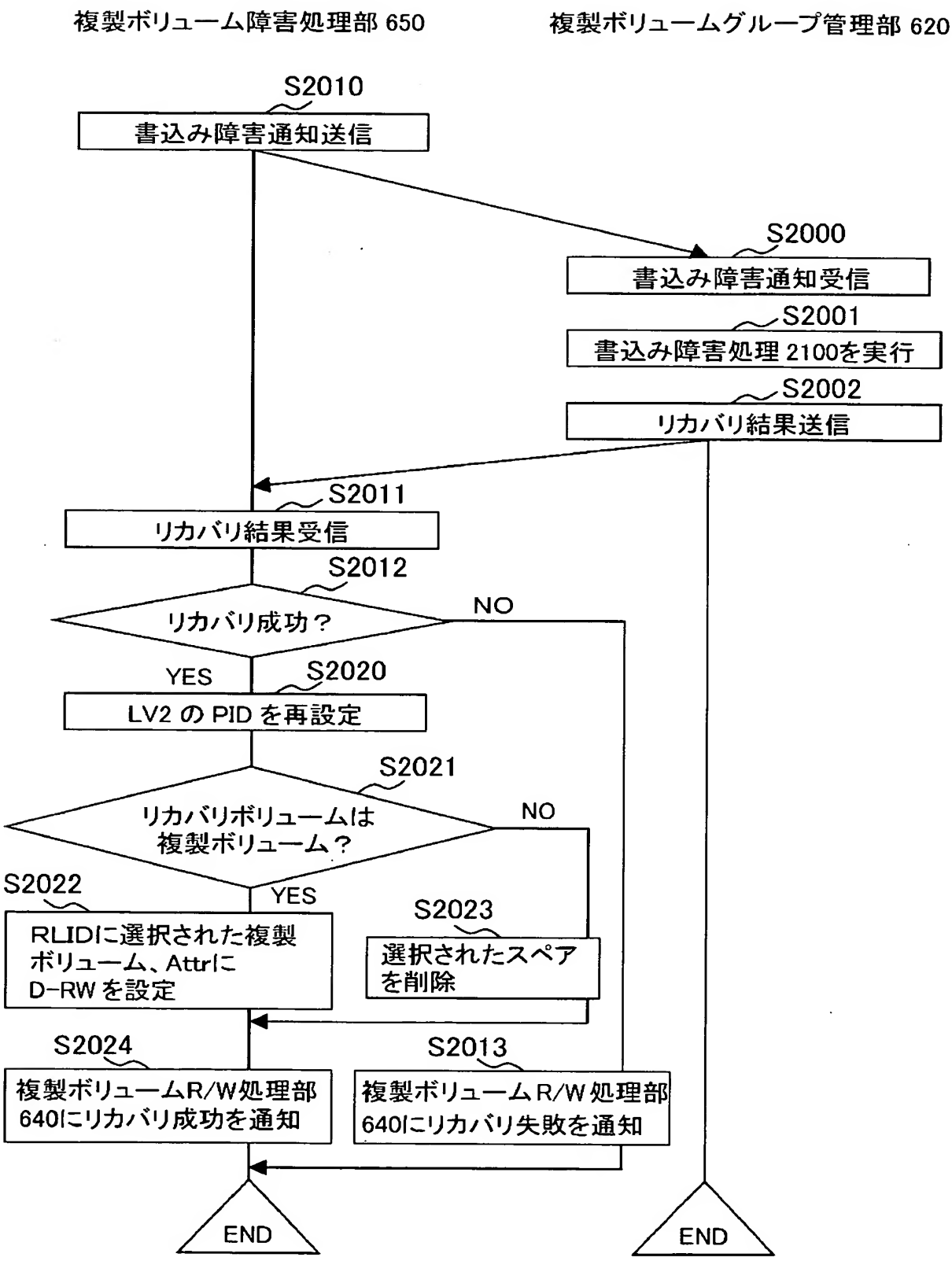

複製ボリューム読出し/書込み処理部(書込み処理)
S1800
RW 複製ボリューム(LV2)への書込み要求受信
書込み要求
wbid
Offset
書込みデータ
S1801
Attr は D-RW?
YES
S1802
分離書込み処理実行
NO
S1810
通常書込み
S1811
書込み障害発生？
NO
YES
S1820
書込み障害を複製ボリューム障害処理部 800 に通知
S1812
Freq を1増やし、書込みリプライ(成功)送信
S1821
YES
リカバリ成功？
NO
S1822
書込みリプライ(失敗)送信
END

【図１９】

複製ボリューム読出し/書込み処理部（分離書込み処理）

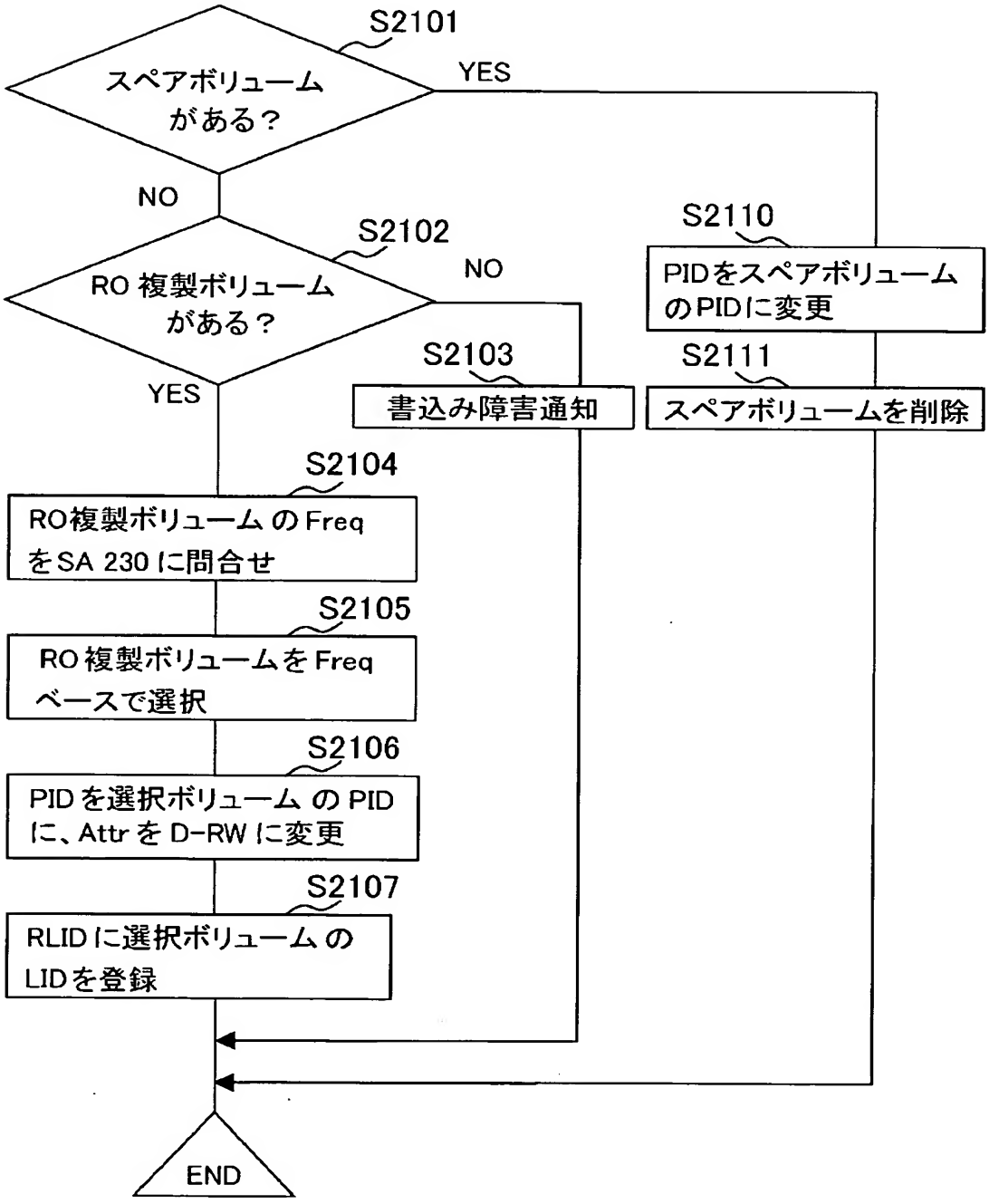


【図２０】

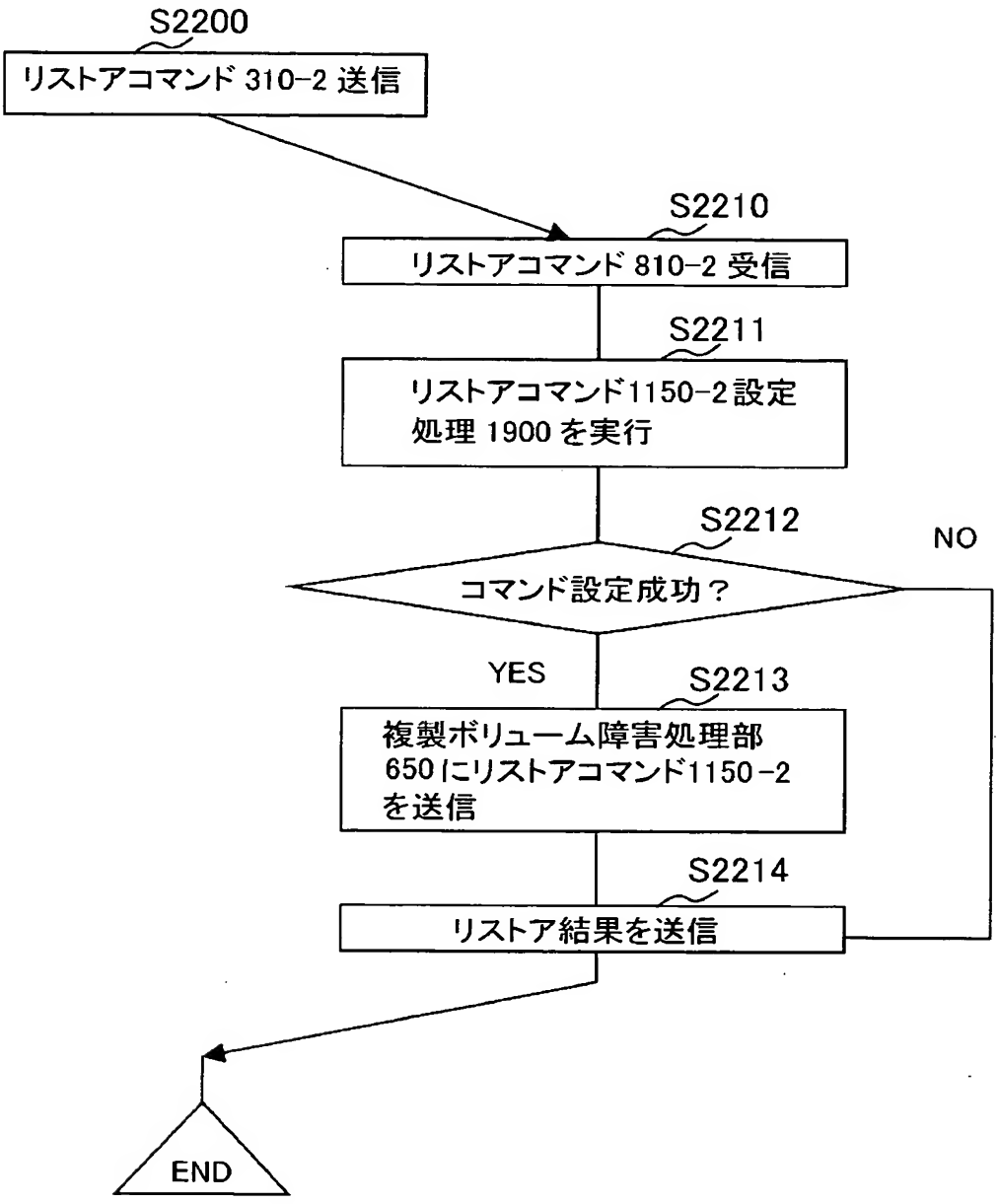

複製ボリューム障害処理部 650
複製ボリュームグループ管理部 620
S2010
書込み障害通知送信
S2000
書込み障害通知受信
S2001
書込み障害処理 2100を実行
S2002
リカバリ結果送信
S2011
リカバリ結果受信
S2012
リカバリ成功？
NO
YES
S2020
LV2 の PID を再設定
S2021
リカバリボリュームは
複製ボリューム？
NO
YES
S2022
RLIDに選択された複製
ボリューム、Attrに
D-RW を設定
S2023
選択されたスペア
を削除
S2024
複製ボリュームR/W処理部
640にリカバリ成功を通知
S2013
複製ボリュームR/W処理部
640にリカバリ失敗を通知
END
END

【図２１】

ボリューム書込み障害処理

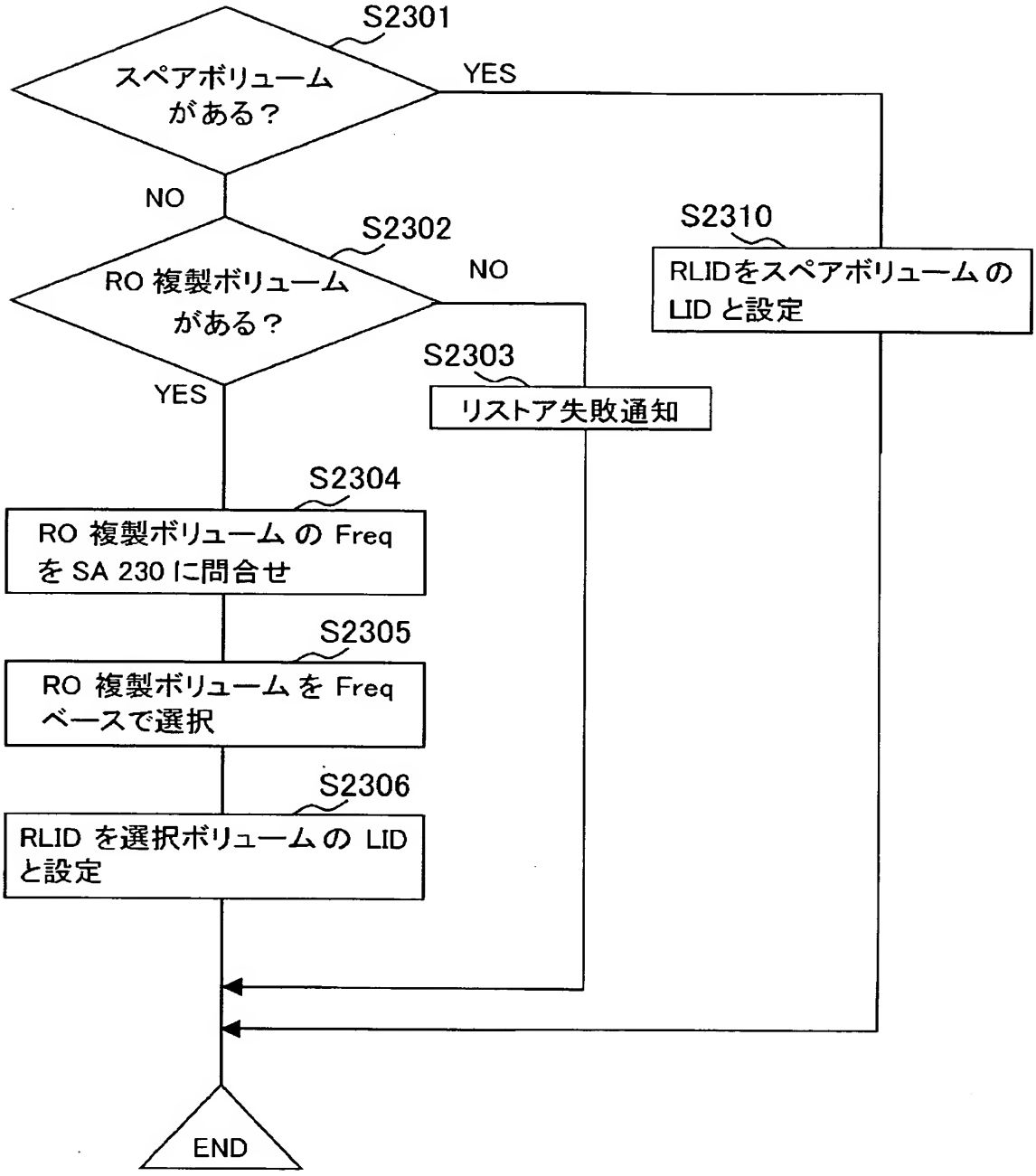


【図２２】

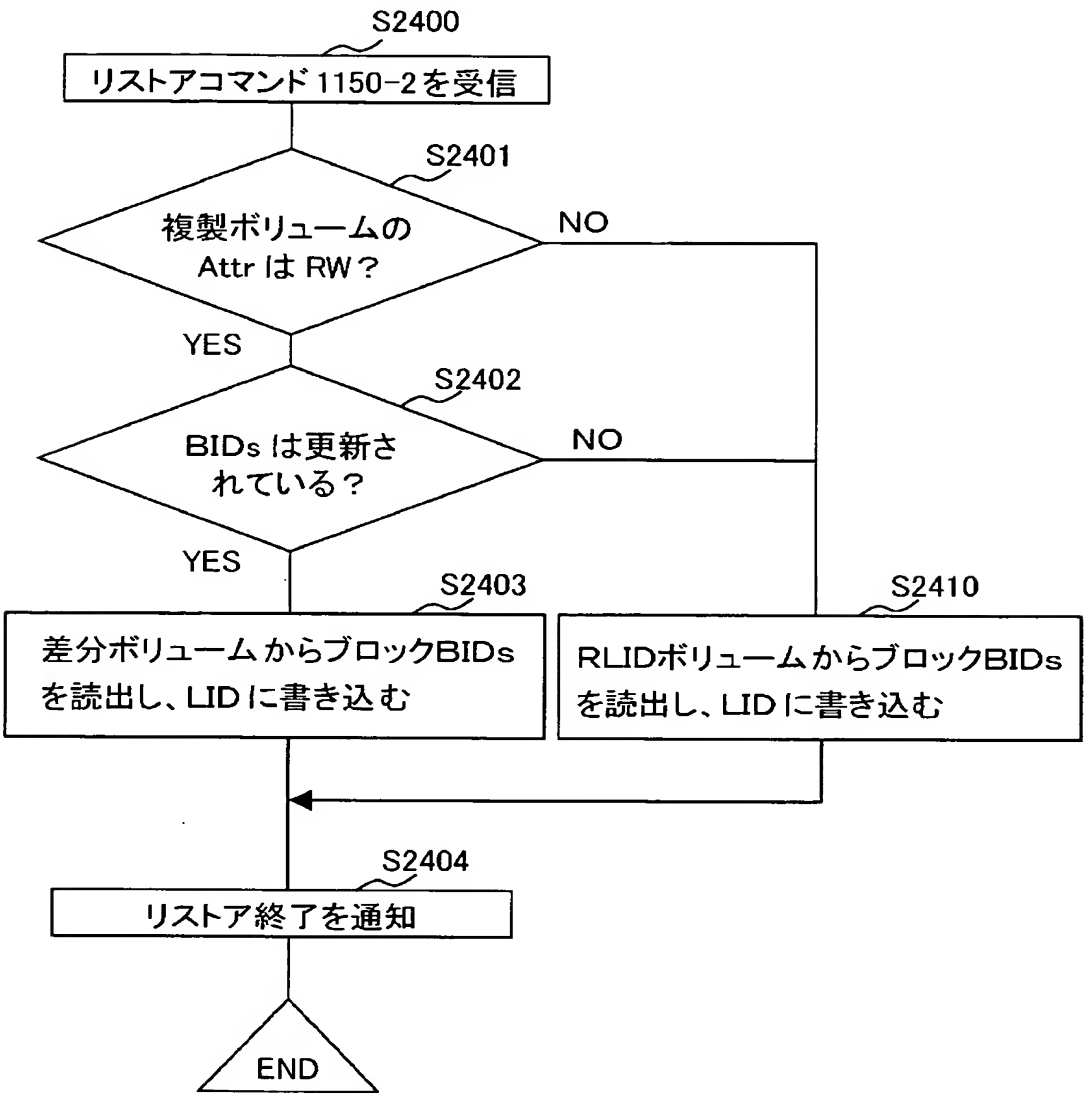

複数ボリュームグループ設定部 610
複数ボリュームグループ管理部 620
S2200
リストアコマンド 310-2 送信
S2210
リストアコマンド 810-2 受信
S2211
リストアコマンド1150-2設定処理 1900 を実行
S2212
コマンド設定成功？
NO
YES
S2213
複製ボリューム障害処理部650にリストアコマンド1150-2を送信
S2214
リストア結果を送信
END

【図２３】

リストアコマンド設定処理



【図２４】


複製ボリューム障害処理部
S2400
リストアコマンド1150-2を受信
S2401
複製ボリュームの
Attr は RW？
NO
YES
S2402
BIDs は更新さ
れている？
NO
YES
S2403
差分ボリュームからブロックBIDs
を読出し、LID に書き込む
S2410
RLIDボリュームからブロックBIDs
を読出し、LID に書き込む
S2404
リストア終了を通知
END

【書類名】要約書
【要約】
【課題】複製ボリュームについて可用性を確保する。
【解決手段】データの記憶領域を供給する記憶装置と、外部からのアクセス要求を受け付けて、前記アクセス要求に応じて前記記憶領域に対するデータの読み出し／書き込みを行うアクセス処理部と、前記記憶領域を用いて構成される論理的な記憶領域（論理ボリューム）を管理する論理ボリューム管理部と、本番系の業務に適用されている前記論理ボリュームである本番ボリュームと、前記本番ボリュームに書き込まれるデータの複製が書き込まれる前記論理ボリュームである複製ボリュームとを管理する複製ボリューム管理部と、
前記複製ボリュームの一つに障害が生じている場合に、当該複製ボリュームとは異なる他の複製ボリュームに書き込まれているデータを用いて、前記障害の内容に応じた方法により、当該複製ボリュームを復元する複製ボリューム復元部と、を備えるデータＩ／Ｏ装置を提供する。
【選択図】　図1

特願２００３－３４３４７８

出　願　人　履　歴　情　報

識別番号　　［０００００５１０８］

１．変更年月日　　１９９０年　８月３１日
［変更理由］　　新規登録
住　所　　東京都千代田区神田駿河台４丁目６番地
氏　名　　株式会社日立製作所